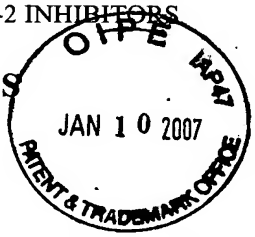
IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN RE APPLICATION OF: ANDREW J. DANNENBERG ET AL.
SERIAL NO.: 10/614,795
FILED: July 9, 2003
FOR: MULTI-FUNCTIONAL COX-2 INHIBITORS

GROUP ART UNIT: 1614
EXAMINER: Lezah Roberts
ATTY. REFERENCE: DANN3009/ESS



COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

Transmitted herewith is a communication/amendment in the above-identified application.

☒ Small entity status under *37 CFR 1.9* and *1.27* is claimed.

☐ No additional fee is required.

The fee, if any, has been calculated as shown below:

Fee Basis	Number of Claims After Amendment	Highest Number Previously Paid For	Extra Claims	Small Entity	Full Fee
Total Claims		- [1]	= [3]	× $ 25 =	× $ 50 =
Independent Claims		- [2]	= [3]	× $100 =	× $ 200 =
☐ First Presentation of Proper Multiple Dependent Claim				+ $180 =	+ $360 =
			TOTAL		

[1] If less than 20 enter **20**. [2] If less than 3 enter **3**. [3] If less than 0 enter **0**.

☐ Please charge my **Deposit Account Number 02-0200** in the amount of $_____. A duplicate copy of this sheet is attached.

☒ A check in the amount of **$ 250.00** is attached. (Check No. 50892)

☒ The Commissioner is hereby authorized to charge any additional fees associated with this communication, including fees due under *37 CFR 1.16* and *37 CFR 1.17* or credit any overpayment to **Deposit Account Number 02-0200**. A duplicate copy of this sheet is attached.

☒ Also enclosed is/are: Appeal Brief.

23364
Customer Number
Phone: (703) 683-0500

DATE: January 10, 2007

Respectfully submitted,

Eric S. Spector
Attorney for Applicant
Registration Number: 22,495

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of	)
ANDREW J. DANNENBERG, et al.	) Group Art Unit: 1614
Patent Application No. 10/614,795	) Examiner: Lezah Roberts
Filed: July 9, 2003	) Confirmation No. 8535
For: MULTI-FUNCTIONAL COX-2 INHIBITORS	)

APPEAL BRIEF PURSUANT TO 37 C.F.R. 41.37

Honorable Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

A Notice of Appeal and Petition for a three month Extension of Time were filed on December 12, 2006. Thus this Appeal Brief is due by February 12, 2007.

Compliance with 37 C.F.R. 41.37 follows:

37 C.F.R. 41.37(c)(1)(i)
Real Party in Interest

The real party in interest is Cornell Research Foundation, Inc.

37 C.F.R. 41.37(c)(1)(ii)
Related Appeals and Interferences

There are no related appeals and interferences.

37 C.F.R. 41.37(c)(1)(iii)
Status of Claims

Claims 6-11 are finally rejected and are the appealed claims.
Claims 1-5 have been canceled.

37 C.F.R. 41.37(c)(1)(iv)
Status of Amendments Subsequent to Final Action

There were no amendments subsequent to final action.

37 C.F.R. 41.37(c)(1)(v)
Summary of the Claimed Invention for Each Independent Claim Involved in the Appeal

The sole independent claim involved in the appeal is claim 6.

The claim is directed to a method for screening a selective inhibitor of COX-2 for likelihood of success in treating a patient having or at risk for cancer, Alzheimer's disease or atherosclerosis.

That the method is for screening a selective inhibitor of COX-2 is indicated in the application as filed at page 1, lines 18-19. That the method is directed to likelihood of success in treatment is set forth at page 11, lines 19-21. That the method is directed to likelihood of success in treatment of a patient having or at

risk for cancer, Alzheimer's disease or atherosclerosis is indicated by reference at page 11, line 19 of the application, to success in treatment in the second embodiment and at page 11, lines 22-26 which indicates that the second embodiment is directed to treating a patient having or at risk for cancer, Alzheimer's disease or atherosclerosis.

Seven tests are enumerated in claim 6. These are set forth in the application as filed at page 3, line 26 – page 4, line 5.

37 C.F.R. 41.37(c)(1)(vi)

Grounds of the Rejection to be Reviewed on Appeal

There are no rejections based on prior art.

The only two rejections are both based on 35 U.S.C. 112, first paragraph.

In the first of the two rejections under 35 U.S.C. 112, first paragraph, claims 6-11 are rejected on the basis that claim 6 violates the enablement requirement of 35 U.S.C. 112, first paragraph.

The rejection pro forma applies the Wands factors but doesn't provide a useful list of violations of the enablement requirement.

In the response of August 15, 2006, the undersigned suggested positions that the PTO is asserting and the Advisory Action of October 17, 2006, only modifies one and suggests no others. The only positions the PTO is asserting, appear to be the following (denoted items (1) – (6)):

(1) It is difficult to predict whether a drug will be effective to treat cancer (relying on Gura and Johnson);

(2) The claims are directed to screening for "likelihood of success" in treating disease and likelihood of success is not defined;

(3) The tests recited in the claims are not sufficient for developing drugs for treating all cancers;

(4) The tests of claim 6 are not enabled for likelihood of success in treating Alzheimer's disease (AD) because AD cannot be diagnosed without autopsy and because evaluation of treatment results is not

possible because of variation of progression of disease from patient to patient (relying on MedicalNet.com);

(5) Data obtained from animal models in development of drugs for treating Alzheimer's disease is not conclusive (relying on Firuzi) when it is not clear how the tests relate to Alzheimer's disease;

(6) The fact that each of the enumerated tests are directed to inflammatory activities is irrelevant because the claims are not directed to likelihood of success in treating inflammation associated with the listed disorders.

In the second of the rejections under 35 U.S.C. 112, first paragraph, claims 6-11 are rejected on the basis that claim 6 violates the written description requirement of 35 U.S.C. 112, first paragraph. The final Office Action explains "The success measurement for each of the diseases now in claim 6 is not disclosed." The Advisory Action indicates the reference to the second embodiment in the phrase "successful for the treatment of or in the second embodiment" at page 11, line 19, doesn't mean success in treating a patient with cancer, Alzheimer's disease or atherosclerosis, even though the second embodiment is directed to treating a patient having or at risk for cancer, atherosclerosis or Alzheimer's disease.

37 C.F.R. 41.37(c)(1)(vii)
ARGUMENT

Rejection Under 35 U.S.C. 112 (second paragraph) – Lack of Enablement

The final rejection takes the position "[t]hat the claims are not enabled because of the screening test [sic] are not related to the disclosed diseases" (see page 3 of the final action). The final rejection then applies the Wands factors to support this conclusion.

The final rejection is defective because the conclusion stated is irrelevant to enablement.

For enablement, the issue rather is whether the specification teaches how to practice (that is, carry out, the invention) without undue experimentation. (see In

re Wands, 8 U.S.P.Q.2d 1400, 1402 (Fed. Cir. 1988) which defines the test for enablement as whether the specification is written to enable those skilled in the art to practice the invention without undue experimentation. See also Ex parte Forman, 230 U.S.P.Q. 546, 547 (Bd. of Pat. App. and Int., 1986), relied on by Wands for the Wands factors, which states: "The ultimate question . . . is whether or not the specification contains a sufficiently explicit disclosure to enable one having ordinary skill in the relevant field to practice the invention claimed therein without the exercise of undue experimentation.

The specific issue in Wands involved whether monoclonal antibodies necessary to practice the immunoassay method claimed were enabled without undue experimentation when practice involved screening negative hybridomas to find those that produced the desired antibodies.

In Forman a question was whether mutant strains of S. typhis necessary for an oral vaccine were enabled when there was a lack of guidance leading to predictable results for obtaining mutant S. typhis.

The instant case differs from the specific issues in Wands and Forman because no issue has been raised that the materials used in the testing are not definite or not described in literature and/or are not articles of commerce. Thus, the specific issues present in Wands and Forman are not present here.

The instant case clearly meets what Wands and Forman state as the general issue, i.e., the specification here contains a detailed exemplified disclosure of how to practice the claimed invention. The method of the claims here is quite straightforward and involves testing to determine meeting at least two of seven criteria (denoted (a), (b), (c), (d), (e), (f) and (g) in claim 6). The tests are described in great detail at pages 4-10 of the application as filed and are used in Working Examples I and II of the application as filed. As indicated above, the specific issues in Wands and Forman are not present here as no issue has been raised that the agents used in the testing are not readily obtained.

Moreover, even if the conclusion used as a basis of the rejection is considered appropriate (and as indicated above, it is not), the positions on which the final rejection and Advisory Action rely to support their conclusion are defective.

Consider firstly that the burden is on the PTO to show that the claimed invention doesn't work. Casting doubt is not enough. See Ex parte Reese 40

U.S.P.Q.2d 1221 (Pat. Off. Bd. App. Int. 1996); In re Dinh-Hguyen, 181 U.S.P.Q. 46, 47 (C.C.P.A. 1974) and In re Gardiner, 177 U.S.P.Q. 396, 397 (C.C.P.A. 1973).

We turn now to the positions which the PTO has asserted as being relevant. These are listed above in 37 C.F.R. 41.37(a)(1)(vi).

Several of these positions are really directed to the same issue, namely why should the claimed invention work. Why should the test work for all cancers (item (3))? Why should the drug determined work for Alzheimer's disease (items (4), and (5))? Why should the tests work to screen for success in treating a patient having or at risk for cancer, Alzheimer's disease or atherosclerosis when the tests are for inflammatory activity (item (6))? The answer to all these questions is that it is recognized by those skilled in the art and it is well known, that inflammation is an inciting event in the pathogenesis (development of disease) of cancer, Alzheimer's disease and atherosclerosis. As indicated above, if the PTO wishes to contest this, the burden is on the PTO to show that this is not so. Confirmation is readily obtained by a search in Google or PubMed for inflammation crossed with pathogenesis of cancer, (14,250 hits in PubMed), atherosclerosis (2,657 hits in PubMed) and Alzheimer's disease (501 hits in PubMed).

In regard to item (4), a point raised is how one can tell success for Alzheimer's disease when variation of progression from patient to patient varies. The issue is not a legitimate issue. There are recognized drug treatments for Alzheimer's disease, e.g. Cogirex, Aricept, Excelon, estrogen, ibuprofen, naprosyn. If the PTO wishes to contest this, the burden is on it to show that there are no recognized drug treatments for Alzheimer's disease. The fact that there are recognized drugs for treatment of Alzheimer's disease, proves that efficacy can be observed.

Moreover, item (3) is irrelevant to claim 7 (colon cancer), claim 8 (Alzheimer's disease), claim 9 (atherosclerosis), claim 10 (oral premalignant lesion of the tongue) and claim 11 (cervical intraepithelial neoplasia).

Moreover, items (4) and (5) are irrelevant to claim 7 (colon cancer), claim 9 (atherosclerosis), claim 10 (oral premalignant lesion of the tongue) and claim 11 (cervical intraepithelial neoplasia).

We turn now to item (1), that it is difficult to predict whether a drug will be effective for cancer. In response, it is noted that the tests are screening tests, i.e., a starting point for progression to further testing. Evidence is presented in a response dated October 15 that in vitro screening can lead to successful clinical results for anti-cancer drugs. See Steinbach, G., et al., The New England Journal of Medicine 324 (26), 1946 - 1952 (2000); Swain, S., The New England Journal of Medicine 353 (26), 2807 – 2809 (2005) and Voskoglou-Nomikos, T., et al., Clinical Cancer Research (9), 4227 – 4239 (2003), copies filed with the response of October 15 and included in the evidence Appendix hereto. This rebuts item (1).

That leaves item (2), that the claims are directed to screening for "likelihood of success" in treating disease and likelihood of success is not defined. It is appellant's position that this would be understood by those skilled in the art as suitable for selection for the next step of testing. The PTO has not disagreed with this position and has only replied that the testing cannot indicate selection suitable for all cancers. This is answered above where it is pointed out that the tests are for inflammatory activities and inflammation is an inciting event in the pathogenesis of cancer as well as in Alzheimer's disease and atherosclerosis.

Thus the Appellants have shown that the instant case meets the general criteria set forth in Wands and Forman and therefore meets the description requirement and additionally that the specific positions asserted by the PTO are irrelevant and rebutted.

Rejection Under 35 C.F.R. 112 (first paragraph) – Lack of Written Description

Appellants have taken the position that written description is present in the application as filed at page 11, lines 17-26.

Originally the position by the PTO in rebuttal was: "The success measurement for each of the diseases now in claim 6 is not disclosed." But that is not recited in the claims. Rather, passing of specific tests is.

The Advisory Action apparently superseded the above with:

> Although Applicant points out the area of support, there appears to be lack of support for screening compounds to treat Alzheimer's, cancer or atherosclerosis. The screening methods are recited to test a candidate for the likelihood of

> success but not the likelihood for success for a specific disease. The diseases are mentioned for the second embodiment for a method to treating these diseases, not a method for screening an agent. Therefore the rejection is maintained.

The undersigned doesn't understand this position. The portion of the sentence at page 11 being referred to apparently is "being successful for the treatment of and in the second embodiment herein." That sentence means being successful for the treatment of the second embodiment herein and being successful in the second embodiment here. As indicated at page 11, the second embodiment is directed to a method of treating a patient having or at risk for cancer, Alzheimer's disease or atherosclerosis. Making substitution of what the second embodiment is, in the quoted sentence, gives -- being successful for the treatment of a patient having or at risk for cancer, Alzheimer's disease or atherosclerosis and being successful in a method for treating a patient having or at risk for cancer, Alzheimer's disease or atherosclerosis --.

In view of the above, it is submitted that the PTO position on 35 U.S.C. 112, first paragraph, is egregiously defective.

The Appendixes

Claims, evidence and proceedings appendixes are attached.

The Fee

The fee required by 37 C.F.R. 41.37(a)(2) and 37 C.F.R. 41.20(b)(2), is enclosed.

Request for Reversal

Reversal of the rejection and allowance is requested.

Respectfully submitted,
BACON & THOMAS, PLLC

By: _____
Eric S. Spector
Registration No. 22,495

BACON & THOMAS, PLLC
Customer 23364
625 Slaters Lane - 4th Floor
Alexandria, VA 22314-1176
Telephone: (703) 683-0500
Facsimile: (703) 683-1080

DANN3009/ESS
CRF D-2756

Date: January 10, 2007

37 C.F.R. 41.37(c)(1)(viii)
Claims Appendix of Claims Involved in Appeal

6. method for screening a selective inhibitor of COX-2 for likelihood of success in treating a patient having or at risk for cancer, Alzheimer's disease or atherosclerosis, comprising testing for at least two of

(a) causing increase in PPRE luciferase activity by at least 100% as manifested by at least doubling of luciferase activity based on data that have been normalized with β-galactosidase activity; (b) causing at least 50% decrease in level of or 50% downregulation of expression of Class I family of receptors tyrosine kinase; (c) causing at least 50% downregulation of expression of cyclin D1; (d) causing at least 50% downregulation of expression of HPV16 oncoproteins E6 and E7; (e) causing at least 50% increase in expression of PTEN, (f) causing at least 50% inhibition of tcf/lef/-catenin-mediated promoter activation; and (g) causing at least 50% increase in level of Nrf-2; the more of (a), (b), (c), (d), (e), (f) and (g) being met, the greater the likelihood of success.

7. The method of claim 6 which is for screening a selective inhibitor of COX-2 for likelihood of success in treating a patient with colon cancer.

8. The method of claim 6 which is for screening a selective inhibitor of COX-2 for likelihood of success in treating a patient with Alzheimer's disease.

9. The method of claim 6 which is for screening a selective inhibitor of COX-2 for likelihood of success in treating a patient with atherosclerosis.

10. The method of claim 6 which is for screening a selective inhibitor of COX-2 for likelihood of success in treating a patient with an oral premalignant lesion of the tongue.

11. The method of claim 6 which is for screening a selective inhibitor of COX-2 for likelihood of success in treating cervical intraepithelial neoplasia.

37 C.F.R. 41.37(c)(1)(ix)
Evidence Appendix

The attached evidence relied on herein was submitted to the PTO with the response of August 15, 2006 and was entered by the Advisory Action of October 17, 2006:

(1) Steinbach, G., et al, New England Journal of Medicine, 324 (26), 1946 – 1952 (2000);

(2) Swain, S., The New England Journal of Medicine 353 (26), 2807 – 2809 (2005); and

(3) Voskoglou-Nomikos, T., et al., Clinical Cancer Research (9), 4227 – 4239 (2003).

37 C.F.R. 41.37(c)(1)(x)

Related Proceedings Appendix

No related proceedings were identified pursuant to paragraph (c)(1)(ii).

The New England Journal of Medicine

THE EFFECT OF CELECOXIB, A CYCLOOXYGENASE-2 INHIBITOR, IN FAMILIAL ADENOMATOUS POLYPOSIS

GIDEON STEINBACH, M.D., PH.D., PATRICK M. LYNCH, M.D., J.D., ROBIN K.S. PHILLIPS, M.B., B.S., MARINA H. WALLACE, M.B., B.S., ERNEST HAWK, M.D., M.P.H., GARY B. GORDON, M.D., PH.D., NAOKI WAKABAYASHI, M.D., PH.D., BRIAN SAUNDERS, M.D., YU SHEN, PH.D., TAKASHI FUJIMURA, M.D., LI-KUO SU, PH.D., AND BERNARD LEVIN, M.D.

ABSTRACT

Background Patients with familial adenomatous polyposis have a nearly 100 percent risk of colorectal cancer. In this disease, the chemopreventive effects of nonsteroidal antiinflammatory drugs may be related to their inhibition of cyclooxygenase-2.

Methods We studied the effect of celecoxib, a selective cyclooxygenase-2 inhibitor, on colorectal polyps in patients with familial adenomatous polyposis. In a double-blind, placebo-controlled study, we randomly assigned 77 patients to treatment with celecoxib (100 or 400 mg twice daily) or placebo for six months. Patients underwent endoscopy at the beginning and end of the study. We determined the number and size of polyps from photographs and videotapes; the response to treatment was expressed as the mean percent change from base line.

Results At base line, the mean (±SD) number of polyps in focal areas where polyps were counted was 15.5±13.4 in the 15 patients assigned to placebo, 11.5±8.5 in the 32 patients assigned to 100 mg of celecoxib twice a day, and 12.3±8.2 in the 30 patients assigned to 400 mg of celecoxib twice a day ($P=0.66$ for the comparison among groups). After six months, the patients receiving 400 mg of celecoxib twice a day had a 28.0 percent reduction in the mean number of colorectal polyps ($P=0.003$ for the comparison with placebo) and a 30.7 percent reduction in the polyp burden (the sum of polyp diameters) ($P=0.001$), as compared with reductions of 4.5 and 4.9 percent, respectively, in the placebo group. The improvement in the extent of colorectal polyposis in the group receiving 400 mg twice a day was confirmed by a panel of endoscopists who reviewed the videotapes. The reductions in the group receiving 100 mg of celecoxib twice a day were 11.9 percent ($P=0.33$ for the comparison with placebo) and 14.6 percent ($P=0.09$), respectively. The incidence of adverse events was similar among the groups.

Conclusions In patients with familial adenomatous polyposis, six months of twice-daily treatment with 400 mg of celecoxib, a cyclooxygenase-2 inhibitor, leads to a significant reduction in the number of colorectal polyps. (N Engl J Med 2000;342:1946-52.)

HUMAN colon cancer develops in a stepwise fashion from normal mucosa to adenomatous polyps to carcinoma. Mutation in the adenomatous polyposis coli (*APC*) gene commonly occurs early in the development of sporadic adenomas.[1] Patients with familial adenomatous polyposis have an inherited germ-line *APC* mutation[2] that results in hundreds of adenomatous polyps and a nearly 100 percent risk of colon cancer. Management includes prophylactic proctocolectomy, or colectomy followed by sigmoidoscopic surveillance and rectal polypectomy. Because the adenoma-to-carcinoma sequence in familial adenomatous polyposis resembles sporadic colon carcinogenesis,[1] studies of familial adenomatous polyposis may contribute to the prevention of sporadic adenomas and colon cancer.

Nonsteroidal antiinflammatory drugs (NSAIDs) reduce the incidence of carcinogen-induced colon tumors in rodents.[3,4] NSAIDs are associated with a reduced incidence of and mortality from sporadic adenoma and colon cancer in epidemiologic studies.[5-8] In early clinical studies[9,10] and small, randomized, placebo-controlled trials,[11-13] sulindac caused the regression of colorectal adenomas in patients with familial adenomatous polyposis. However, the gastrointestinal toxicity associated with conventional NSAIDs may limit their long-term use for cancer prevention.[14]

NSAIDs are inhibitors of the cyclooxygenase enzyme family, which catalyzes the metabolism of arachidonic acid to prostaglandins, prostacyclin, and thromboxanes. The cyclooxygenase-1 isoform is constitutively expressed in most tissues, where it medi-

From the University of Texas M.D. Anderson Cancer Center, Houston (G.S., P.M.L., N.W., Y.S., T.F., L.-K.S., B.L.); the Imperial Cancer Research Fund, St. Mark's Hospital, London (R.K.S.P., M.H.W., B.S.); the National Cancer Institute, Bethesda, Md. (E.H.); and G.D. Searle, Skokie, Ill. (G.B.G.). Address reprint requests to Dr. Steinbach at the University of Texas M.D. Anderson Cancer Center, Division of Cancer Prevention, Box 203, 1515 Holcombe Blvd., Houston, TX 77030, or at gsteinb@aol.com.

Other authors were Louis Godio, Ph.D. (G.D. Searle, Skokie, Ill.), Sherri Patterson, B.A. (M.D. Anderson Cancer Center, Houston), Miguel A. Rodriguez-Bigas, M.D. (Roswell Park Cancer Institute, Buffalo, N.Y.), Susan L. Jester, M.S. (G.D. Searle), Karen L. King, M.S. (G.D. Searle), Marta Schumacher, M.B.A. (National Cancer Institute, Bethesda, Md.), James Abbruzzese, M.D. (M.D. Anderson Cancer Center), Raymond N. DuBois, M.D., Ph.D. (Vanderbilt University Medical Center, Nashville), Walter N. Hittelman, Ph.D. (M.D. Anderson Cancer Center), Stuart Zimmerman, Ph.D. (M.D. Anderson Cancer Center), Jeffrey W. Sherman, M.D. (G.D. Searle), and Gary Kelloff, M.D. (National Cancer Institute).

ates physiologic functions such as gastric mucosal cytoprotection and regulation of platelet aggregation. Its inhibition may account for many of the common side effects of NSAIDs, including gastric ulceration and gastrointestinal hemorrhage.[14,15] The cyclooxygenase-2 isoform is induced in response to cytokines and growth factors and is expressed in inflammatory disease, premalignant lesions (such as colorectal adenomas), and colon cancer.[16-18] Its inhibition has not been associated with gastric ulceration.[15,19-21] However, the long-term effects of selective cyclooxygenase-2 inhibitors as compared with those of traditional NSAIDs remain to be determined.[22] Experimental evidence supports the concept that the chemopreventive effects of NSAIDs may be due at least in part to inhibition of cyclooxygenase-2.[23,24] Hence, selective inhibition of cyclooxygenase-2 offers a potential pharmacologic strategy for the prevention of colorectal adenomas.

To determine whether inhibition of cyclooxygenase-2 could reduce the extent of polyposis in patients with familial adenomatous polyposis, we studied the effect of celecoxib, an agent that selectively inhibits cyclooxygenase-2.[21]

METHODS

Patients

Patients with familial adenomatous polyposis who were 18 to 65 years of age, who had not had their entire colorectum removed, and who had five or more polyps 2 mm or more in diameter that could be assessed endoscopically, were eligible. Exclusion criteria included a history of colectomy within the previous 12 months or colectomy anticipated within 8 months after randomization; use of NSAIDs or aspirin three or more times a week within 6 months of randomization or one or two times a week within 3 months of randomization; or abnormal results of serum laboratory tests (complete blood count and liver-function and renal-function tests).

The study was approved by the institutional review board of the University of Texas M.D. Anderson Cancer Center and the ethics committee of St. Mark's Hospital, London. Written informed consent was obtained from all patients.

Study Design

The study was randomized, double-blinded, and placebo-controlled. It was conducted between December 1996 and December 1998 at the M.D. Anderson Cancer Center in Houston and St. Mark's Hospital in London. One hundred eight patients who were eligible for screening underwent endoscopy; 29 had insufficient polyps for inclusion in the study, and 2 required colectomy for advanced disease (a rectal cancer and a large sessile adenoma). According to the protocol, 75 patients were initially randomly assigned in a 2:2:1 ratio to receive celecoxib (Celebrex, G.D. Searle, Skokie, Ill.), either 100 mg twice daily or 400 mg twice daily, or an identical-appearing placebo orally for six months. The placebo contained 250 mg of lactose. Two additional patients were assigned to the group receiving 100 mg of celecoxib twice daily after two patients were withdrawn because of noncompliance. The study drug and matching placebo were provided by G.D. Searle.

The six-month duration of the study and the end point of adenoma regression were based on previous trials of sulindac that demonstrated an effect on polyp regression within six months of treatment.[9-13] A clinical trial aimed at the prevention of carcinoma, on the other hand, would require many years of study and therefore was not considered feasible for the initial testing of the efficacy of a drug. Evaluations at base line and month 6 included a history taking, physical examination, and endoscopy, with biopsies of the intact or residual colorectum, stomach, and duodenum. Testing for *APC* gene mutations was performed at base line.[25]

Compliance was monitored by means of pill counts and review of diaries completed by the patients. Safety was monitored with a comprehensive symptom questionnaire administered by telephone at two-to-four-week intervals that elicited information on adverse events and by clinical laboratory evaluations at base line and at one, three, and six months. Adverse events were graded in accordance with the National Cancer Institute Common Toxicity Criteria.[26]

Endoscopy

At the base-line endoscopy, an India-ink tattoo was placed in the rectum, colon, or both near a small area with a high density of polyps. The base-line and six-month endoscopic examinations were videotaped, and a series of photographs was taken with the tattoo, appendix, or ileocecal valve positioned centrally and peripherally. These photographs were used for quantitative measurements of the number and size of polyps. Polyps for biopsy were taken from areas that were not photographed for scoring.

Enumeration and Measurement of Polyps

To ascertain that the same area was scored at base line and at month 6, polyps were counted in pairs of photographs. One investigator, other than the endoscopist, who did not know the treatment, performed the scoring. Videotapes were used to resolve ambiguities and confirm polyp counts. The diameter of a polyp was measured with the aid of a standardized endoscopic ruler or biopsy forceps included in the photographic field to serve as a scale. Because in patients with familial adenomatous polyposis the colon is studded with microscopic and poorly visible lesions, only distinct polyps at least 2 mm in diameter were counted.

A qualitative assessment of the total extent of colorectal polyposis was conducted by each of five endoscopists experienced in the management of familial adenomatous polyposis (two from each of the study centers and one from a nonparticipating polyposis center) during joint videotape-review sessions. The first of each pair of videos (obtained at base line and month 6) was scored as the same as, better than, or worse than the second, without the endoscopists' being aware of the temporal sequence or treatment group. A score of "better" or "worse" indicated that there was a clear difference in the total extent of polyp involvement. To avoid bias, videotapes of three colorectal regions (cecum and ascending colon; transverse, descending, and sigmoid colon; and rectum) were assessed separately without the endoscopists' being aware of whether the segments came from the same patient.

Statistical Analysis

All 77 randomly assigned patients were included in the intention-to-treat analysis of toxicity and polyp number, size, and burden. Analysis of the endoscopic videotape assessments was performed in the patients for whom the requisite videotapes were available.

The quantitative response variables were the percent change from base line in polyp number and polyp burden, defined as the sum of the polyp diameters. The percent change in each patient was calculated on the basis of the photographs at the tattoo, appendix, and ileocecal valve, and the mean change was then calculated for each study group. Efficacy was evaluated by comparing the mean percent change from base line in each treatment group with that in the placebo group by the Wilcoxon rank-sum test.

Whether treatment affected the polyp count at six months was also analyzed in a multivariate linear regression model with adjustment for base-line covariates. Two variables indicating the treatment (100 or 400 mg twice a day) were included in the model, and the other base-line covariates were the number of polyps, sex, age, study site, and surgical status (whether the patient had previously

undergone colectomy). We employed a logarithmic transformation of both the base-line and the final polyp-count values to eliminate the skewness in that distribution.

In the qualitative assessment of response, based on review of the endoscopic videotapes, each segment was assigned a score of 1 for better, 0 for same, or −1 for worse, and the mean of the five physicians' scores for each treatment group was compared with that for the placebo group with use of the Wilcoxon rank-sum test. The response of each videotaped colorectal segment (cecum and ascending colon; transverse, descending, and sigmoid colon; and rectum) was analyzed separately. In addition, the response of the total colorectum, defined for each patient as the mean score for all colorectal segments assessed, was analyzed.

Adverse events, including those with an onset within 30 days after the end of treatment, were coded according to World Health Organization Adverse Reaction Terminology and graded for severity with the National Cancer Institute Common Toxicity Criteria.[26] Clinical laboratory data were compared between treatment groups by one-way analysis of variance applied to the change from base line to month 1, month 3, month 6, or early termination.

The Kruskal–Wallis test was used to compare base-line continuous variables among the three treatment groups, and the chi-square test or Fisher's exact test was used to examine associations between nominal variables. All tests were two-sided, and a P value of less than 0.05 was considered to indicate statistical significance.[27] No interim analyses were performed.

RESULTS

Patients

Seventy-seven patients were enrolled: 36 at the M.D. Anderson Cancer Center and 41 at St. Mark's Hospital. The treatment groups were similar with regard to race or ethnic group, sex ratio, surgical status, and number of polyps, but they differed in age: the group assigned to 400 mg of celecoxib twice a day was younger (33.1 years) than the group assigned to 100 mg of celecoxib twice a day (38.6 years) and the placebo group (39.9 years) (Table 1). Sixty-six patients had an identified *APC* mutation, and two additional patients had relatives with known *APC* mutations. Seventy-two of the 77 patients completed the treatment. More than 90 percent of the patients who completed the study took at least 80 percent of the study drug. At base line, the placebo group had a mean (±SD) of 15.5±13.4 polyps, the group assigned to 100 mg of celecoxib twice a day had a mean of 11.5±8.5 polyps, and the group assigned to 400 mg of celecoxib twice a day had a mean of 12.3±8.2 polyps in the focal areas where polyps were counted (P=0.66 for the comparison among groups).

TABLE 1. BASE-LINE CHARACTERISTICS OF THE PATIENTS WITH FAMILIAL ADENOMATOUS POLYPOSIS.*

CHARACTERISTIC	PLACEBO (N=15)	100 mg OF CELECOXIB TWICE DAILY (N=32)	400 mg OF CELECOXIB TWICE DAILY (N=30)	P VALUE
Age — yr	39.9±11.3	38.6±10.0	33.1±10.9	0.04†
Sex — no. (%)				0.84‡
Male	9 (60)	17 (53)	18 (60)	
Female	6 (40)	15 (47)	12 (40)	
Race or ethnic group — no. (%)				0.87§
Black	0	1 (3)	1 (3)	
White	15 (100)	29 (91)	26 (87)	
Hispanic	0	2 (6)	3 (10)	
Height — cm	171.5±7.7	169.9±9.7	169.1±11.6	0.74†
Weight — kg	74.6±16.4	74.4±12.7	71.1±15.4	0.39†
Surgical status — no. (%)				0.45‡
Intact colon	5 (33)	8 (25)	12 (40)	
Colectomy	10 (67)	24 (75)	18 (60)	
No. of polyps	15.5±13.4	11.5±8.5	12.3±8.2	0.66†
Polyp size — mm	2.9±0.5	2.9±0.7	2.9±0.6	0.63†
Polyp burden — mm¶	44.7±36.5	34.8±28.1	37.6±29.4	0.65†

*Plus–minus values are means ±SD.

†The P value was calculated by the Kruskal–Wallis test.

‡The P value was calculated by the chi-square test.

§The P value was calculated by Fisher's exact test.

¶The polyp burden was calculated as the sum of the polyp diameters.

Response to Treatment

Treatment with 400 mg of celecoxib twice daily for six months was associated with a significant reduction from base line in the number of colorectal polyps as compared with the placebo group (28.0 percent vs. 4.5 percent, P=0.003) (Table 2 and Fig. 1). The group receiving 100 mg of celecoxib twice daily had a reduction of 11.9 percent as compared with 4.5 percent in the placebo group (P=0.33). Multivariate linear regression analysis confirmed that 400 mg of celecoxib twice daily reduced the number of colorectal polyps (P=0.005) after adjustment for age, sex, surgical status (colectomy vs. intact colon), number of polyps at base line, and investigational institution.

A reduction of 25 percent or more in the mean number of colorectal polyps was seen in 53 percent of the patients in the group receiving 400 mg of celecoxib twice daily (P=0.003 for the comparison with placebo), 31 percent of the patients in the group receiving 100 mg of celecoxib twice daily (P=0.08), and 7 percent of patients in the placebo group. Intention-to-treat analysis of the specific response of rectal polyps as distinct from colonic polyps showed a mean reduction in the number of rectal polyps of 22.5 percent (P=0.01 for the comparison with the placebo group) in the group receiving 400 mg of celecoxib twice daily and of 3.4 percent (P=0.52 for the comparison with the placebo group) in the group receiving 100 mg of celecoxib twice daily, as compared with a mean increase of 3.1 percent in the placebo group (Table 2).

Whereas the number of polyps was quantified in designated small areas adjacent to a tattoo or anatomical landmark, the full extent of colorectal polyposis was assessed qualitatively from videotapes of complete anatomical segments of the colorectum by a panel of five endoscopists. The videotapes showed that in the group receiving 400 mg of celecoxib twice daily, sig-

TABLE 2. PERCENT CHANGE FROM BASE LINE IN THE MEAN NUMBER OF POLYPS AND COLORECTAL POLYP BURDEN IN PATIENTS WITH FAMILIAL ADENOMATOUS POLYPOSIS TREATED WITH PLACEBO OR CELECOXIB FOR SIX MONTHS.*

VARIABLE	PLACEBO (N=15)	100 mg OF CELECOXIB TWICE DAILY (N=32)	400 mg OF CELECOXIB TWICE DAILY (N=30)
Percent change in no. of colorectal polyps	−4.5±16.4	−11.9±30.3	−28.0±24.0
P value		0.33	0.003
Percent change in colorectal polyp burden†	−4.9±17.3	−14.6±31.7	−30.7±25.7
P value		0.09	0.001
Percent change in no. of rectal polyps‡	+3.1±31.1	−3.4±35.0	−22.5±26.0
P value		0.52	0.01

*Plus–minus values are means ±SD. P values are based on the two-sample Wilcoxon statistic for the comparison of celecoxib with placebo, in the intention-to-treat analysis. Negative numbers indicate decreases, and positive numbers increases.

†The colorectal polyp burden was calculated as the sum of the polyp diameters.

‡Seven subjects had no rectal polyps at base line or on final evaluation. These subjects are considered to have had 0 percent change.

nificant improvement in polyposis occurred in the rectum (P=0.01), in the ascending colon and cecum (P=0.02), and in the transverse, descending, and sigmoid colon (P=0.003) (Table 3). The corresponding changes in the group receiving 100 mg of celecoxib twice daily were not significant, but there was a trend toward a dose response in the rectum (P=0.07) and in the ascending colon and cecum (P=0.10). The combined assessments from all the videotapes of the colon and rectum showed a consistent improvement in the group receiving 400 mg of celecoxib twice daily (P<0.001) as well as in the group receiving 100 mg twice daily (P=0.03).

To estimate changes in polyp area, the polyp burden was calculated as the sum of the polyp diameters. The average decreases in polyp burden were 30.7 percent for the group receiving 400 mg of celecoxib twice daily, 14.6 percent for the group receiving 100 mg of celecoxib twice daily, and 4.9 percent for the placebo group (P=0.001 for the comparison of 400 mg of celecoxib twice daily and placebo) (Table 2).

Safety

Both doses of celecoxib were well tolerated. Sixty-eight percent of the patients in the placebo group, 56 percent of the patients receiving 100 mg of celecoxib twice daily, and 57 percent of the patients receiving 400 mg of celecoxib twice daily reported one or more adverse events of grade 2 or higher according to the National Cancer Institute Common Toxicity Criteria.[26] Of these, the most commonly reported (by at least 10 percent of patients in each treatment group) were diarrhea (placebo, 13 percent; 100 mg of celecoxib twice daily, 19 percent; 400 mg of celecoxib twice daily, 13 percent) and abdominal pain (placebo, 13 percent; 100 mg of celecoxib twice daily, 3 percent; 400 mg of celecoxib twice daily, 7 percent). There were no significant differences in the incidence of any adverse event between the celecoxib groups and the placebo group. In addition to two patients with-

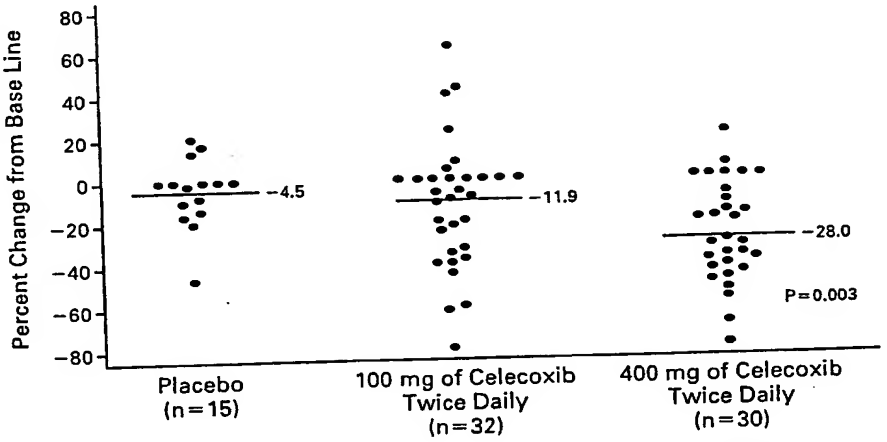


Figure 1. Percent Change from Base Line in the Number of Colorectal Polyps in 77 Patients with Familial Adenomatous Polyposis Who Were Treated with Placebo or Celecoxib (100 mg Twice a Day or 400 mg Twice a Day) for Six Months.

A decrease from base line represents disease regression, and an increase represents disease progression. The horizontal lines show the mean changes. The P value is for the comparison with the placebo group.

TABLE 3. CHANGE IN COLORECTAL POLYPOSIS BASED ON REVIEW OF ENDOSCOPIC VIDEOTAPES IN PATIENTS WITH FAMILIAL ADENOMATOUS POLYPOSIS TREATED WITH PLACEBO OR CELECOXIB FOR SIX MONTHS.*

COLORECTAL SEGMENT	PLACEBO	100 mg OF CELECOXIB TWICE DAILY†	400 mg OF CELECOXIB TWICE DAILY‡
Rectum			
No. of patients	15	29	29
Score	−0.1±0.3	0.2±0.5	0.3±0.4
P value		0.07	0.01
Transverse, descending, and sigmoid colon			
No. of patients	6	3	10
Score	−0.2±0.2	−0.1±0.1	0.4±0.4
P value		0.33	0.003
Cecum and ascending colon			
No. of patients	5	7	10
Score	−0.2±0.4	0.4±0.6	0.5±0.4
P value		0.10	0.02
Total colorectum§			
No. of patients	15	29	29
Score	−0.07±0.26	0.13±0.22	0.33±0.32
P value		0.03	<0.001

*The base-line and six-month evaluations were compared by a panel of endoscopists experienced in the management of familial adenomatous polyposis; these endoscopists assigned scores for anatomical segments at six months as follows: −1 indicated "worse," 0 "no change," and 1 "better." Plus–minus values are means ±SD of the scores for each group. P values are based on the two-sample Wilcoxon statistic for the comparison of celecoxib with placebo, in the analysis of patients for whom the respective videotapes were available.

†Videotapes were not available for three patients.

‡Videotapes were not available for one patient.

§The score for the total colorectum is the mean of the separate assessments of the transverse, descending, and sigmoid colon; the cecum and ascending colon; and the rectum.

drawn for noncompliance, three patients did not complete the study for the following reasons: suicide in a patient in the group receiving 100 mg twice daily with a history of psychiatric disorder and a previous suicide attempt, acute allergic reaction in a patient in the group receiving 400 mg twice daily with a history of allergies, and dyspepsia in a patient in the group receiving 400 mg twice daily. There were no significant alterations in mean laboratory-test values. No ulceration was seen on follow-up esophagogastroduodenoscopy in any patient, including the patient who withdrew because of dyspepsia.

After the study was completed, patients were not offered continuation of treatment with the study drug because the efficacy of the drug was not known until the results were analyzed. Three patients (one from each study group) are known to have undergone colectomy since the completion of the study.

DISCUSSION

In a six-month study, we found that treatment with a cyclooxygenase-2 inhibitor, celecoxib, at a dose of 400 mg twice daily was associated with significant regression of colorectal adenomas in patients with familial adenomatous polyposis. Significant regression was not associated with the dose of 100 mg twice daily. These clinical findings are consistent with other evidence that cyclooxgenase-2 has a role in colonic tumorigenesis and that selective inhibition of cyclooxygenase-2 may help control this process.[23]

Regression of adenomas was seen in the rectum as well as in the left and right sides of the colon. Age and whether or not the patient had undergone colectomy did not affect the results. Nonetheless, our six-month study leaves many important questions unanswered. These include whether prolonged treatment with a medication such as celecoxib can replace, delay the need for, or limit the anatomical extent of proctocolectomy, and whether such treatment can inhibit progression to carcinoma. Our findings suggest, however, that celecoxib could serve as an adjunct to current management by suppressing polyp formation in patients with residual rectum after colectomy and in patients with an intact colon who are awaiting colectomy.

Sulindac, a nonselective cyclooxygenase inhibitor, was previously reported to cause complete or nearly complete regression of rectal adenomas in uncontrolled studies,[9,10,28] and in a small, placebo-controlled, drug-crossover trial of patients with familial adenomatous polyposis.[11] Regression of rectal adenomas, though of lesser magnitude, was reported in two subsequent placebo-controlled studies, by Giardiello et al.[12] and Nugent et al.[13] In the former study, 12 patients treated with sulindac showed maximal improvement by month 6 of the nine-month study. In contrast to earlier reports, no patient had a complete remission, and the clinical effect was considered insufficient to eliminate the need for colectomy in patients with established polyposis. Rapid recurrence of adenomas was also observed three to four months after discontinuation of drug therapy.[11,12] Evidence of long-term efficacy of sulindac is still lacking, and there have been case reports of tumor progression in patients receiving sulindac.[29] Because of differences in patients' characteristics and in study methods, differences in findings among these studies cannot be critically assessed. Long-term studies, as well as direct comparisons of selective and nonselective cyclooxygenase inhibition, could further define the relative clinical benefits of these drugs.

A key question is whether the inhibitory effect of NSAIDs on colon carcinogenesis is mediated by inhibition of either cyclooxygenase-1 or cyclooxygenase-2, or both, or by inhibition of other cellular targets of NSAIDs. Several lines of evidence indicate that cyclooxygenase-2 mediates this process, although noncyclooxygenase pathways may also be involved.[23,30-32] Cyclooxygenase-2 is up-regulated in colonic neoplasms, including adenomas and carcinomas in humans and rodents, and in early adenomas in mice with

germ-line *APC* mutations.[17,24,33] Selective cyclooxygenase-2 inhibition reduces the incidence of carcinogen-induced colonic aberrant crypt foci and carcinomas in rats, as well as the incidence of adenomas in mice with germ-line *APC* mutations.[24,34,35] There is also direct genetic evidence that the cyclooxygenase-2 gene contributed to the development of adenomas in a mouse model of familial adenomatous polyposis, in which knockout of the cyclooxygenase-2 gene greatly reduced the number of intestinal adenomas.[24] Such studies support the concept that the antineoplastic effects of NSAIDs are attributable, at least in part, to inhibition of cyclooxygenase-2.

The specific cellular pathways responsible for the effects of cyclooxygenase-2 on tumorigenesis are under study. Current evidence indicates that cyclooxygenase-2 mediates mitogenic growth factor signaling and down-regulates apoptosis, thus promoting tumor growth.[36-38] The induction of apoptosis by selective inhibition of cyclooxygenase-2 is relevant to familial adenomatous polyposis, in which apoptosis is considered to be attenuated.[39]

Preclinical studies have established the role of cyclooxygenase-2 in colon tumorigenesis and suggested a role for cyclooxygenase-2 inhibition in the prevention of human cancer. Our findings support the application of this strategy to studies of the prevention of colorectal tumors in other populations at risk, including persons with sporadic adenomatous polyps in whom cellular tumorigenesis resembles familial adenomatous polyposis. The role of cyclooxygenase-2 inhibition in preventing adenomas in adolescents with preclinical familial adenomatous polyposis remains to be studied.

Supported by a contract with the National Cancer Institute (NO1 CN-65118), a Cancer Center Support Grant (CA-16672), and contracts with Searle Pharmaceuticals (protocol IQ4-96-02-001).

Drs. Steinbach, Phillips, Wallace, and Levin have served as consultants to G.D. Searle.

We are indebted to Diane Gravel, R.N., for patient care, Ms. Jill Sawyer for genetic counseling, and Nancy Matteer, B.S., for assistance with laboratory tests at the M.D. Anderson Cancer Center; to Kay Neale, Nicola Baxter, and Elizabeth Avery for patient care at the Polyposis Registry of St. Mark's Hospital; to Drs. Randolph H. Bailey, Monica M. Bertagnolli, and Henry T. Lynch for referral of patients; and to the patients and their families for participating in and contributing to this demanding study.

REFERENCES

1. Powell SM, Zilz N, Beazer-Barclay Y, et al. APC mutations occur early during colorectal tumorigenesis. Nature 1992;359:235-7.
2. Kinzler KW, Nilbert MC, Su LK, et al. Identification of FAP locus genes from chromosome 5q21. Science 1991;253:661-5.
3. Pollard M, Luckert PH. Indomethacin treatment of rats with dimethylhydrazine-induced intestinal tumors. Cancer Treat Rep 1980;64:1323-7.
4. Reddy BS, Bao CV, Rivenson A, Kelloff G. Inhibitory effect of aspirin on azoxymethane-induced colon carcinogenesis in F344 rats. Carcinogenesis 1993;14:1493-7.
5. Kune GA, Kune S, Watson LF. Colorectal cancer risk, chronic illnesses, operations, and medications: case control results from the Melbourne Colorectal Cancer Study. Cancer Res 1988;48:4399-404.
6. Thun MJ, Namboodiri MM, Heath CW Jr. Aspirin use and reduced risk of fatal colon cancer. N Engl J Med 1991;325:1593-6.
7. Rosenberg L, Palmer JR, Zauber AG, Warshauer ME, Stolley PD, Shapiro S. A hypothesis: nonsteroidal anti-inflammatory drugs reduced the incidence of large-bowel cancer. J Natl Cancer Inst 1991;83:355-8.
8. Giovannucci E, Rimm EB, Stampfer MJ, Colditz GA, Ascherio A, Willett WC. Aspirin use and the risk for colorectal cancer and adenoma in male health professionals. Ann Intern Med 1994;121:241-6.
9. Waddell WR, Loughry RW. Sulindac for polyposis of the colon. J Surg Oncol 1983;24:83-7.
10. Rigau J, Pique JM, Rubio E, Planas R, Tarrech JM, Bordas JM. Effects of long-term sulindac therapy on colonic polyposis. Ann Intern Med 1991;115:952-4.
11. Labayle D, Fischer D, Vielh P, et al. Sulindac causes regression of rectal polyps in familial adenomatous polyposis. Gastroenterology 1991;101:635-9.
12. Giardiello FM, Hamilton SR, Krush AJ, et al. Treatment of colonic and rectal adenomas with sulindac in familial adenomatous polyposis. N Engl J Med 1993;328:1313-6.
13. Nugent KP, Farmer KC, Spigelman AD, Williams CB, Phillips RK. Randomized controlled trial of the effect of sulindac on duodenal and rectal polyposis and cell proliferation in patients with familial adenomatous polyposis. Br J Surg 1993;80:1618-9.
14. Wolfe MM, Lichtenstein DR, Singh G. Gastrointestinal toxicity of nonsteroidal antiinflammatory drugs. N Engl J Med 1999;340:1888-99. [Erratum, N Engl J Med 1999;341:548.]
15. Vane JR, Botting RM. Mechanism of action of anti-inflammatory drugs. Scand J Rheumatol Suppl 1996;102:9-21.
16. Gierse JK, Hauser SD, Creely DP, et al. Expression and selective inhibition of the constitutive and inducible forms of human cyclo-oxygenase. Biochem J 1995;305:479-84.
17. Eberhart CE, Coffey RJ, Radhika A, Giardiello FM, Ferrenbach S, DuBois RN. Up-regulation of cyclooxygenase 2 gene expression in human colorectal adenomas and adenocarcinomas. Gastroenterology 1994;107:1183-8.
18. Kargman SL, O'Neill GP, Vickers PJ, Evans JF, Mancini JA, Jothy S. Expression of prostaglandin G/H synthase-1 and -2 protein in human colon cancer. Cancer Res 1995;55:2556-9.
19. Lipsky PE, Isakson PC. Outcome of specific COX-2 inhibition in rheumatoid arthritis. J Rheumatol 1997;24:Suppl 49:9-14.
20. Simon LS, Lanza FL, Lipsky PE, et al. Preliminary study of the safety and efficacy of SC-58635, a novel cyclooxygenase 2 inhibitor: efficacy and safety in two placebo-controlled trials in osteoarthritis and rheumatoid arthritis, and studies of gastrointestinal and platelet effects. Arthritis Rheum 1998;41:1591-602.
21. Geis GS. Update on clinical developments with celecoxib, a new specific COX-2 inhibitor: what can we expect? J Rheumatol 1999;26:Suppl 56:31-6.
22. Physicians' desk reference. 54th ed. Montvale, N.J.: Medical Economics, 2000.
23. Taketo MM. Cyclooxygenase-2 inhibitors in tumorigenesis. J Natl Cancer Inst 1998;90:1609-20.
24. Oshima M, Dinchuk JE, Kargman SL, et al. Suppression of intestinal polyposis in Apc delta716 knockout mice by inhibition of cyclooxygenase 2 (COX-2). Cell 1996;87:803-9.
25. Kinzler KW, Nilbert MC, Su LK, et al. Identification of FAP locus genes from chromosome 5q21. Science 1991;253:661-5.
26. Cancer Therapy Evaluation Program. Common toxicity criteria. Bethesda, Md.: National Cancer Institute, March 1998.
27. Altman DG. Practical statistics for medical research. London: Chapman & Hall, 1990.
28. Winde G, Schmid KW, Schlegel W, Fischer R, Osswald H, Bunte H. Complete reversion and prevention of rectal adenomas in colectomized patients with familial adenomatous polyposis by rectal low-dose sulindac maintenance treatment: advantages of a low-dose nonsteroidal anti-inflammatory drug regimen in reversing adenomas exceeding 33 months. Dis Colon Rectum 1995;38:813-30.
29. Niv Y, Fraser GM. Adenocarcinoma in the rectal segment in familial polyposis coli is not prevented by sulindac therapy. Gastroenterology 1994;107:854-7.
30. Kopp E, Ghosh S. Inhibition of NF-kappa B by sodium salicylate and aspirin. Science 1994;265:956-9.
31. He T-C, Chan TA, Vogelstein B, Kinzler KW. PPARδ is an APC-regulated target of nonsteroidal anti-inflammatory drugs. Cell 1999;99:335-45.
32. Piazza GA, Alberts DS, Hixson LJ, et al. Sulindac sulfone inhibits azoxymethane-induced colon carcinogenesis in rats without reducing prostaglandin levels. Cancer Res 1997;57:2909-15.
33. Boolbol SK, Dannenberg AJ, Chadburn A, et al. Cyclooxygenase-2

overexpression and tumor formation are blocked by sulindac in a murine model of familial adenomatous polyposis. Cancer Res 1996;56:2556-60.
34. Reddy BS, Rao CV, Seibert K. Evaluation of cyclooxygenase-2 inhibitor for potential chemopreventive properties in colon carcinogenesis. Cancer Res 1996;56:4566-9.
35. Kawamori T, Rao CV, Seibert K, Reddy BS. Chemopreventive activity of celecoxib, a specific cyclooxygenase-2 inhibitor, against colon carcinogenesis. Cancer Res 1998;58:409-12.
36. Coffey RJ, Hawkey CJ, Damstrup L, et al. Epidermal growth factor receptor activation induces nuclear targeting of cyclooxygenase-2, basolateral release of prostaglandins, and mitogenesis in polarizing colon cancer cells. Proc Natl Acad Sci U S A 1997;94:657-62.
37. Vadlamudi R, Mandal M, Adam L, Steinbach G, Mendelsohn J, Kumar R. Regulation of cyclooxygenase-2 pathway by HER2 receptor. Oncogene 1999;18:305-14.
38. Tsujii M, DuBois RN. Alterations in cellular adhesion and apoptosis in epithelial cells overexpressing prostaglandin endoperoxide synthase 2. Cell 1995;83:493-501.
39. Pasricha PJ, Bedi A, O'Connor K, et al. The effects of sulindac on colorectal proliferation and apoptosis in familial adenomatous polyposis. Gastroenterology 1995;109:994-8.

Evidence Appendix
Document (2)

The NEW ENGLAND JOURNAL of MEDICINE

EDITORIALS

Aromatase Inhibitors — A Triumph of Translational Oncology

Sandra M. Swain, M.D.

Great strides have been made in the diagnosis and treatment of early-stage breast cancer, thanks to advances in molecular medicine, interdisciplinary treatment, and rapid electronic communication. Hormonal therapy, the first and most successful targeted therapy for breast cancer, has saved many thousands of lives. Moreover, screening and adjuvant (postoperative) therapy have increased survival among women with breast cancer.[1,2] The improvement in survival can be attributed to both adjuvant tamoxifen therapy and adjuvant chemotherapy and has been found in all subgroups of patients regardless of the presence or absence of tumor cells in draining lymph nodes, including women who are premenopausal, those who are postmenopausal, those with estrogen-receptor–negative tumors, and those with estrogen-receptor–positive tumors. Experts are now in the process of classifying breast cancer, which actually consists of a heterogeneous group of cancers, into multiple categories. It is essential to define each subgroup precisely and to delineate distinct characteristics and targets that will lead to tailored therapies that are better than the ones we have now.

In this issue of the *Journal*, the Breast International Group (BIG) 1-98 Collaborative Group reports on a randomized comparison of letrozole, an aromatase inhibitor, with tamoxifen as adjuvant therapy for postmenopausal women with early-stage breast cancer. Their findings validate the results of previous studies showing that aromatase inhibitors were more efficacious than tamoxifen in such women.[3] The BIG 1-98 Collaborative Group found a reduction in the incidence of relapse of 3.4 percentage points at five years in the letrozole group, as compared with the tamoxifen group, after a median follow-up of 25.8 months. The incidence of both distant recurrence and contralateral breast cancers was reduced. The benefit was greatest in patients who had also received chemotherapy, who did not receive radiotherapy, and who had positive nodes. Longer follow-up is important to define the benefit of letrozole in patients with node-negative disease. There was no significant difference in survival between the two groups, but at this point, fewer deaths have occurred among women assigned to letrozole.

Five other large trials have also evaluated aromatase inhibitors. The Arimidex, Tamoxifen, Alone or in Combination (ATAC) trial, with a median follow-up of 68 months, found that, as compared with tamoxifen, adjuvant treatment with anastrozole reduced the recurrence rate by 3.7 percentage points in patients with hormone-receptor–positive tumors.[4] The MA.17 trial, in which women first received tamoxifen for five years and then were randomly assigned to receive placebo or letrozole, found that letrozole improved disease-free survival by 4.6 percentage points, after a median follow-up of 30 months, with a survival difference in the node-positive group only.[5] The Intergroup Exemestane Study (IBS), with a median follow-up of 30.6 months, compared 2 to 3 years of tamoxifen followed by 2 to 3 years of exemestane with 5 years of tamoxifen therapy and found that the former regimen increased disease-free survival by 4.7 percentage points.[6] The Italian Anastrozole Trial (ITA), with a median follow-up of 36 months, compared 2 to 3 years of tamoxifen followed by 2 to 3 years of anastrozole with 5 years of tamoxifen and found that sequential treatment re-

2S07

duced recurrent-free survival by 5.8 percentage points.[7] Finally, a combined analysis of data from two prospective, multicenter, randomized trials (the Austrian Breast and Colorectal Cancer Study Group trial 8 plus the Arimidex–Nolvadex study) compared women who received two years of tamoxifen followed by three years of anastrozole with women who were given tamoxifen for five years. After a median follow-up of 28 months, sequential therapy was associated with an event-free survival rate that was 3.1 percentage points higher than the rate associated with tamoxifen alone.[8] These five studies varied with respect to the number of women with hormone-receptor–positive tumors, node-negative tumors, and node-positive tumors and the definition of outcomes. It is clear, however, that these trials, with close to 30,000 participants, consistently demonstrate that treatment with an aromatase inhibitor alone or after tamoxifen treatment is beneficial. The questions that remain are the optimal duration of treatment with an aromatase inhibitor, whether tamoxifen or an aromatase inhibitor should be given first, whether sequential treatment is optimal, which aromatase inhibitor is best, and whether an aromatase inhibitor is beneficial for premenopausal women after ovarian ablation. The decrease in contralateral cancers among women treated with an aromatase inhibitor has important implications for chemoprevention. Ongoing trials should answer each of these questions.

One of the most exciting aspects of the findings of these evaluations of aromatase inhibitors is that an animal model predicted the results. In tumor cells and peripheral tissues in postmenopausal women, estrogen is synthesized by aromatase from androstenedione and testosterone. A mouse model was developed to simulate the hormonal milieu in postmenopausal women and used to investigate the ability of aromatase inhibitors and tamoxifen to hinder the growth of breast-cancer cells.[9] This model predicted a superior clinical outcome with aromatase inhibitors. The same model also predicts that the administration of letrozole alone will be more effective than the sequential administration of tamoxifen and letrozole.[10] Future analyses of the continued follow-up of the BIG 1-98 study, which includes a group randomly assigned to receive letrozole before tamoxifen therapy and a group assigned to receive letrozole after tamoxifen therapy, will answer this important question.

A hypothesis developed from the ATAC study is that estrogen-receptor–positive, progesterone-receptor–negative tumors are more susceptible to anastrozole than tumors that have both types of hormone receptors.[11] Although this hypothesis was not supported by the findings of the BIG 1-98 study, because of the relatively short follow-up and multiple subgroup analyses in the study, the idea also cannot be ruled out. Data that support a differential benefit in patients with progesterone-receptor–negative tumors include the finding that patients with such tumors are likely to have HBR-1–positive or HBR-2–positive breast cancer, positive nodes, tumors with high rates of proliferation and aneuploidy, and lower median levels of estrogen receptors. All these features are typical of an aggressive tumor.[12] Another area of fertile research is the crosstalk between growth factor signaling pathways and the estrogen receptor. This crosstalk may result in tamoxifen resistance by potentiating agonist properties of tamoxifen.

It is clear that unlike tamoxifen, aromatase inhibitors are not associated with an increased risk of thromboembolism or uterine cancer. The incidence of fractures and arthralgias is, however, increased among women taking these inhibitors. Both complications are the result of estrogen deficiency, and they require a thorough evaluation with the aim of limiting these adverse effects. In the BIG 1-98 study, the incidence of serious cardiac events was significantly higher among women given letrozole than among those given tamoxifen. An increase in cardiovascular events among patients receiving an aromatase inhibitor has also been suggested in the IBS and ATAC studies. This finding may be due to a cardioprotective effect of tamoxifen, but whatever the mechanism, the potential for adverse cardiovascular events needs close and careful evaluation.

We have seen a substantial increase in the number of patients with small, node-negative tumors over the past several years. In the future, molecular characterization of individual tumors will assist in determining the metastatic potential of the tumor and its sensitivity to various agents. It is our responsibility as physicians to determine the appropriate adjuvant treatment for patients,

but the choices are increasingly complex. Fortunately, we have the results of large, prospective, well-designed, and well-executed clinical trials, such as BIG 1-98, to facilitate our recommendations. We await longer follow-up from all the studies to enable us to offer patients sound advice regarding the benefits and long-term risks of aromatase inhibitors. Meanwhile, all the evidence points to aromatase inhibitors as critically important for improving the outcome among postmenopausal women with breast cancer who have positive or negative lymph nodes and who are at a substantial risk for recurrent disease.

No potential conflict of interest relevant to this article was reported.

From the Breast Cancer Section, Center for Cancer Research, National Cancer Institute, Bethesda, Md.

1. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005;365:1687-717.
2. Berry DA, Cronin KA, Plevritis SK, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. N Engl J Med 2005;353:1784-92.
3. The Breast International Group (BIG) 1-98 Collaborative Group. A comparison of letrozole and tamoxifen in postmenopausal women with early breast cancer. N Engl J Med 2005;353:2747-57.
4. Howell A, Cuzick J, Baum M, et al. Results of the ATAC (Arimidex, Tamoxifen, Alone or in Combination) trial after completion of 5 years' adjuvant treatment for breast cancer. Lancet 2005;365:60-2.
5. Goss PE, Ingle JN, Martino S, et al. Randomized trial of letrozole following tamoxifen as extended adjuvant therapy in receptor-positive breast cancer: updated findings from NCIC CTG MA.17. J Natl Cancer Inst 2005;97:1262-71.
6. Coombes RC, Hall E, Gibson LJ, et al. A randomized trial of exemestane after two to three years of tamoxifen therapy in postmenopausal women with primary breast cancer. N Engl J Med 2004;350:1081-92. [Erratum, N Engl J Med 2004;351:2641.]
7. Boccardo F, Rubagotti A, Puntoni M, et al. Switching to anastrozole versus continued tamoxifen treatment of early breast cancer: preliminary results of the Italian Tamoxifen Anastrozole Trial. J Clin Oncol 2005;23:5138-47.
8. Jakesz R, Jonat W, Gnant M, et al. Switching of postmenopausal women with endocrine-responsive early breast cancer to anastrozole after 2 years' adjuvant tamoxifen: combined results of ABCSG trial 8 and ARNO 95 trial. Lancet 2005;366:455-62.
9. Yue W, Zhou D, Chen S, Brodie A. A new nude mouse model for postmenopausal breast cancer using MCF-7 cells transfected with human aromatase gene. Cancer Res 1994;54:5092-5.
10. Long BJ, Jelovac D, Handratta V, et al. Therapeutic strategies using the aromatase inhibitor letrozole and tamoxifen in a breast cancer model. J Natl Cancer Inst 2004;96:456-65.
11. Dowsett M, Cuzick J, Wale C, Howell T, Houghton J, Baum M. Retrospective analysis of time to recurrence in the ATAC trial according to hormone receptor status: an hypothesis-generating study. J Clin Oncol 2005;23:7512-7.
12. Arpino G, Weiss H, Lee AV, et al. Estrogen receptor-positive, progesterone receptor-negative breast cancer: association with growth factor receptor expression and tamoxifen resistance. J Natl Cancer Inst 2005;97:1254-61.

Trial Registration Report Card

Jeffrey M. Drazen, M.D., and Alastair J.J. Wood, M.D.

One measure of medical progress is new treatments. The discovery of a novel therapy takes time and money, but more important, it requires the mutual effort of groups that, while they share the common goal of improved treatment, often have fundamentally competing interests. These interests intersect at the clinical trial. Patients who are looking for more effective and safer treatment agree to take part in a clinical trial in the hope that they will benefit from such treatment or that others with similar conditions will benefit later. The company developing the new therapy shares the hope that the trial will be successful, because it wants to market the tested therapy exclusively and profitably for as long as possible before its competitors can launch a similar therapy into the marketplace. These goals, though overlapping, are inevitably in conflict and will generate tension. Such tension has been thrown into sharp relief over the past 15 months by the push for clinical trial registration.

The academic establishment and patients have argued that when patients, motivated by altruism, participate (or even consider participating) in a clinical trial, they are entitled to understand fully all the options available to them in the various trials that are currently recruiting subjects. In addition, their participation in a clinical trial should result in generalizable knowledge that will be available to future patients and investigators to improve patient care. This can happen only when appropriate details of the clinical trial are made available to the public in a timely fashion. The Internet and public registries have made this possible.

Some in industry have argued that to open

CORRECTION

Aromatase Inhibitors — A Triumph of Translational Oncology

Aromatase Inhibitors — A Triumph of Translational Oncology . On page 2807, the sentence that begins five lines from the bottom of the right-hand column should have read, "The Italian Anastrozole Trial (ITA), with a median follow-up of 36 months, compared 2 to 3 years of tamoxifen followed by 2 to 3 years of anastrozole with 5 years of tamoxifen and found that sequential treatment increased recurrent-free survival by 5.8 percentage points," not "reduced recurrent-free survival," as printed.

N Engl J Med 2006;355:533-a

Evidence Appendix Document (3)

Vol. 9, 4227–4239, September 1., 2003

Clinical Predictive Value of the *in Vitro* Cell Line, Human Xenograft, and Mouse Allograft Preclinical Cancer Models[1]

Theodora Voskoglou-Nomikos,[2] Joseph L. Pater, and Lesley Seymour

National Cancer Institute of Canada Clinical Trials Group, Cancer Clinical Trials Division, Cancer Research Institute, Queen's University, Kingston, Ontario, K7L 3N6 Canada

ABSTRACT

Purpose: We looked at the value of three preclinical cancer models, the *in vitro* human cell line, the human xenograft, and the murine allograft, to examine whether they are reliable in predicting clinical utility.

Experimental Design: Thirty-one cytotoxic cancer drugs were selected. Literature was searched for drug activity in Phase II trials, human xenograft, and mouse allografts in breast, non-small cell lung, ovary, and colon cancers. Data from the National Cancer Institute Human Tumor Cell Line Screen were used to calculate drug *in vitro* preclinical activity for each cancer type. Phase II activity *versus* preclinical activity scatter plot and correlation analysis was conducted for each model, by tumor type (disease-oriented approach), using one tumor type as a predictor of overall activity in the other three tumor types combined (compound-oriented approach) and for all four tumor types together.

Results: The *in vitro* cell line model was predictive for non-small cell lung cancer under the disease-oriented approach, for breast and ovarian cancers under the compound-oriented approach, and for all four tumor types together. The mouse allograft model was not predictive. The human xenograft model was not predictive for breast or colon cancers, but was predictive for non-small cell lung and ovarian cancers when panels of xenografts were used.

Conclusions: These results suggest that under the right framework and when panels are used, the *in vitro* cell line and human xenograft models may be useful in predicting the Phase II clinical trial performance of cancer drugs. Murine allograft models, as used in this analysis, appear of limited utility.

INTRODUCTION

Both basic science studies and clinical trials are essential components of the cancer drug discovery process. Potential therapeutics found to be significantly better than no treatment or standard therapies (*i.e.*, active) in preclinical laboratory cancer models or compounds with novel chemotypes and equivalent effectiveness to standard treatments are advanced to confirmatory testing in early (Phase I and II) clinical trials. Considering that RR[3] is a reasonable surrogate end point for survival (required but not sufficient), a favorable RR in Phase II trials advances a drug into additional clinical testing and is considered a prerequisite of drug success in the clinic.

Advancing of a candidate drug from preclinical testing in the laboratory to testing in Phase II clinical trials is based on the assumption that drug activity in cancer models translates into at least some efficacy in human patients, *i.e.*, that cancer laboratory models are clinically predictive. In addition, the relevance of tumor type-specific preclinical results for the corresponding human cancers in the clinic can be viewed through two different approaches: compound-oriented, where a drug is assumed to have potential activity against all human tumor types if it is effective against a single test tumor type, and disease-oriented, where a drug with preclinical activity in a single tumor type would only be expected to be effective in the same tumor type in patients.

Although widely adopted, the above-mentioned assumption and approaches have not been confirmed by studies to date. In addition, all studies aimed to examine the clinical predictive value of laboratory cancer models inevitably suffer from inherent bias because compounds with no activity in preclinical models are generally not advanced to clinical trials.

This work was undertaken to examine the clinical predictive value of three preclinical cancer models that have found wide use: the human *in vitro* cell line; the mouse allograft; and the human xenograft. In these models, tumor volume or life span (*in vivo* mouse models) or cell growth (*in vitro* cell lines) is compared between the treatment group receiving the new drug and a control group (active or inactive control).

The use of preclinical cancer models for selection of potential cancer therapeutics was pioneered by the NCI in the United States in the mid-1950s. The screening strategies used until 1990 were essentially compound oriented and involved a

Received 3/26/03; revised 6/2/03; accepted 6/4/03.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
[1] Supported in part by the National Cancer Institute of Canada, Clinical Trials Group. Presented in part at the 2002 Annual Meeting of the American Society of Clinical Oncology (Abstract 360).
[2] To whom requests for reprints should be addressed, at National Cancer Institute of Canada Clinical Trials Group, Cancer Clinical Trials Division, Cancer Research Institute, Queen's University, 10 Stuart Street, Kingston, Ontario, K7L 3N6 Canada. Phone: (613) 533-6430; Fax: (613) 533-2941; E-mail: dnc mikos@ctg.queensu.ca.

[3] The abbreviations used are: RR, response rate; NCI, National Cancer Institute; NSCLC, non-small cell lung cancer; NSC, National Service Center; T/C%, treated over control tumor volume ratio.

small number of predominantly murine allograft tumors, with emphasis on leukemia (1–7). Several studies from the NCI and others demonstrated that this approach had low clinical predictive value for activity in Phase II trials (5–9) and yielded compounds with selective activity toward human leukemias and lymphomas (10–12). Thus, in 1990, the NCI introduced a disease-oriented *in vitro* Human Tumor Cell Line Screen comprised of 60 cell lines from the most common adult tumors (13–17). The screen was designed so that each tumor type was represented by a panel of cell lines, selected on the basis of different subhistological features, and common drug resistance profiles. It was hoped that this screen would help identify drug leads with high potency and/or selective activity against particular tumor types.

Recently, the NCI examined the correlation between drug activity in Phase II clinical trials and preclinical activity in cancer models (18). Important findings were: (*a*) with the exception of NSCLC, preclinical activity in human xenografts of a particular tumor type did not correlate significantly with Phase II activity in the same type of tumor, (*b*) with the exception of breast and colon histologies, human xenografts did not significantly predict Phase II clinical activity in other cancers types; and (*c*) compounds that were active in at least one-third of all tested human xenografts were likely to have at least some activity in Phase II clinical trials.

Studies examining the clinical predictive value of preclinical cancer models outside the scope of the NCI screening programs have focused on the human xenograft model and have looked predominately into same-tumor correlations (disease-oriented approach). These studies have produced both positive (the model was found clinically predictive) and negative (the model was found to have no clinical predictive value) results in various tumor types (19–27).

Two major criticisms can be made on the overall body of literature concerning the clinical predictive value of preclinical cancer models. First, the vast majority of studies to date, both within and outside the NCI, have based their conclusions on the observation of trends rather than the use of statistical methods. Second, all studies conducted previously have used dichotomous definitions of preclinical and/or clinical activity based on largely invalidated cutoff values of measures of activity: a 20% RR in Phase II clinical trials and (most commonly) a 42% T/C% in human xenografts and mouse allografts.

In addition, two important questions have not been addressed at all by previous studies: the clinical predictive value of the *in vitro* cell line model and the relative clinical usefulness of the different preclinical cancer models in use today (*i.e.*, how different models compare with each other in terms of their ability to identify clinically effective drugs).

Thus, we conducted a study comparing the clinical (Phase II) predictive value of three widely used preclinical laboratory cancer models, the *in vitro* human cell line, the mouse allograft, and the human xenograft. We used quantitative measures of both clinical and preclinical activity and statistical methods. We considered three relevant questions: (*a*) the clinical predictive value of the three models within the same tumor type (disease-oriented approach); (*b*) the clinical predictive value of the three models when one preclinical tumor type is used as a predictor of overall clinical activity in all other tumor types (compound-oriented approach); and (*c*) the clinical predictive value of the three models when overall preclinical and clinical activity in all tumor types combined is considered.

MATERIALS AND METHODS

Study Design

A retrospective, literature-based study was conducted. Data were retrieved from studies published between 1985 and 2000. This period was chosen as one when all three preclinical cancer models of interest to this study were in use and because it was long enough and close enough to the present as to afford data on a relatively large number of recently developed drugs.

The data search was restricted to four of the most common and commonly studied solid tumor types, breast, colorectal, ovarian, and non-small cell lung cancers, to ensure that sufficient data would be available.

The Medline and CancerLit databases were used for the collection of published data. In an attempt to minimize publication bias, both paper publications (peer reviewed) and meeting abstracts (nonpeer reviewed) were used as sources of information. If published data were not available for identified drugs, manufacturers were contacted for unpublished data.

Selection of Drugs

Drugs were identified by searching the Medline and CancerLit databases for compounds that had undergone single agent Phase I clinical trial testing either in 1991 or 1992. Agents with novel targets such as signal transduction or angiogenesis modulators were not included.

This Phase I-based approach to agent identification was used to ensure selection of agents developed within the study time frame of 1985–2000: agents with a published Phase I clinical trial in 1991 or 1992 were expected to have been through preclinical testing between 1985 and 1990 and to have undergone Phase II clinical evaluation by the year 2000. In addition, this approach was adopted to minimize publication bias: publication of Phase I trials is generally less dependent on the observation of favorable tumor responses than publication of Phase II trials or of preclinical cancer model experiments.

Data Collection and Drug Activity

Phase II Clinical Trials. Phase II clinical trials for each drug were identified by searching the Medline and CancerLit databases for scientific papers, reviews, or meeting abstracts. Duplicate publications were discarded. For trials with only abstract information, an additional search by author and/or institution name was conducted in Medline or CancerLit. Scientific papers were used in preference to abstracts, where possible.

Two restrictions were applied. The first was a geographic restriction: to ensure uniform methodology in trial conduct and RR assessment, only Phase II trials conducted in the Americas, Western Europe and Australia were included in the analysis. The second restriction referred to the treatment population and aimed to ensure that uniformly responsive populations of patients would be considered. For breast and ovarian cancer, only Phase II trials that included patients who had received prior chemotherapy for metastatic disease were used, whereas for

NSCLC and colon cancers, the Phase II trials selected included patients who had received no prior chemotherapy.

For each individual Phase II trial the following information was collected: disease site; previous chemotherapy; disease stage; number of patients entered; eligible; evaluable and evaluable for response; number of complete and partial responses; and criteria used for response (standard WHO *versus* other). Trials had to have enrolled a minimum of 14 patients, at least 12 of whom must have been evaluable for response. Completed Phase II trials for which >20% of entered patients were listed as inevaluable for response were considered methodologically unacceptable and were not used. For trials in progress at the time of reporting (meeting abstract format only), the available data were used even if they represented <80% of the enrolled patients, provided that they met the 14-patient criterion. If a trial publication did not specify the previous chemotherapy treatment status of patients, it was not used. Information from Phase I-II trials was used only when the Phase I and II components of the trial were separately conducted and reported. Phase II information was collected regardless of drug dose and route of administration.

For a given drug, in a given cancer type, the activity in a single Phase II clinical trial was recorded as the RR: the number of partial and complete tumor responses over the total number of patients evaluable for response. The number of evaluable rather than eligible patients was used to accommodate information from trials for which final results were not available. In the very few cases where the number of patients evaluable for response was not provided, the number of evaluable patients, the number of eligible patients, or the number of patients entered in the trial (whichever was provided by the investigators) in that priority order was used.

To obtain a drug's overall clinical activity in multiple Phase II trials of patients with the same tumor type, all responses and the collective number of patients evaluable for response were pooled from individual trials to calculate an overall RR. Finally, to get the Phase II activity for any three or four cancer types combined, the individual tumor RRs were averaged.

Human Xenografts and Mouse Allografts. The search strategy for mouse cancer model data were similar to the Phase II process. The only exclusion in this case were results obtained with mouse tumors that were engineered to have special characteristics such as, for example, overexpression of proteins conferring drug resistance.

For each murine allograft or human xenograft, numerical value(s) of activity for drugs of interest was retrieved only if expressed as the treated over control tumor volume ratio (T/C%) or the tumor volume growth inhibition ratio (GI%; and T/C% = 100% − GI%) in the literature sources. In addition, only T/C% values calculated by the formula T/C% = [(RVtreated)/(RVcontrol)] × 100% were collected (where RV = relative volume), whereas T/C% values defined for regressions [T/C% = [(RVtreated (0) − RVtreated(t))/RVtreated (0)] × 100%] were excluded to ensure uniform calculation methods. If the T/C% was not provided but a relative tumor growth curve was given as a figure in a publication, the numerical values for the treatment and control groups provided in this graph were used to calculate the T/C%. Activity reported as all mice cured or 100% complete responses was considered equivalent to and recorded as a T/C% = 0. If no exact T/C% value was given but an interval of values was provided instead (*i.e.*, T/C% >42), a T/C% equal to the interval midpoint value (*i.e.*, a T/C% = 71) was assigned. Finally, where preclinical activity was reported as GI%, it was converted to T/C% by the formula T/C% = 100% − GI%. The activity value for the most effective, nontoxic dose in each schedule was recorded.

Single tumor type preclinical activity of each drug in the murine allograft or human xenograft models was defined as the mean T/C% value from all tested allografts/xenografts of that tumor type. Where the same laboratory had tested a single xenograft/allograft with multiple schedules of the same drug and/or where the same xenograft/allograft had been tested with the same drug by more than one laboratories, T/C% values for a single tumor were obtained by first averaging the same laboratory T/C% values and then the same xenograft T/C% values.

Overall preclinical activity in xenografts/allografts for all four tumor types together was expressed as the average of single tumor mean T/C% values.

***In Vitro* Human Tumor Cell Lines.** The publicly available data from the NCI's Human Tumor Cell Line Screen was used as the information source for the *in vitro* tumor cell line model. Information from the NCI *in vitro* Human Tumor Cell Line Screen was favored because it was a readily available, well-defined, comprehensive, validated, and extensive single source of data. Another important reason was that as an exploratory literature search showed, there was such a wide variation between different investigators in the types of assays used and the nature of cell lines tested that it would have been impossible to comprehensively combine published data from various laboratories.

Acquisition of NCI Human Tumor Cell Line Screen data were done through the internet.[4] Information for each drug was obtained through its NCI code number or NSC number. Such numbers, where available, were identified either from the literature or from a cross-reference of compound names and NSC numbers in the NCI database (also available on the NCI web site).[4]

Testing of compounds in the NCI *in vitro* Human Tumor Cell Line Screen has been described previously (17). Briefly, growth inhibition in cell lines is measured by the GI_{50}, defined as the drug concentration that causes a 50% reduction in cell number in test plates relative to control plates. For every drug entering the screen, a concentration range comprised of five, 10-fold dilutions is tested in each of a group of 60–80 cell lines. The optical densities between treated and control plates, as resulting from the sulforhodamine B assay, are used to construct a dose-response curve for each cell line in the screen, leading to the calculation of a GI_{50} in every case by interpolation. In the case of compounds with low (*i.e.*, the highest concentration tested causes <50% growth inhibition) or high (*i.e.*, the lowest concentration tested causes >50% growth inhibition) potency where interpolation is not possible, the highest and lowest concentrations, respectively, in the tested drug concentration

[4] Internet address: http://www.dtp.ncl.nih.gov/docs/cancer/searches/cancer_open_compounds.html.

range are recorded as the approximated GI_{50}s. GI_{50}s are then converted to their Log_{10} values and the overall mean $Log_{10}GI_{50}$ across all cell lines in the screen is calculated. Finally, the results are displayed by a bar graph called the mean graph (28). This graph lists all of the cell lines and their corresponding $Log_{10}GI_{50}$s and relates the magnitude of every individual cell line $Log_{10}GI_{50}$ to the mean $Log_{10}GI_{50}$ across all of the cell lines by a bar to the right (more sensitive than average) or to the left (less sensitive than average) of a vertical line. The experiment is repeated several times for each concentration range. In cases where mean graphs are based on mostly approximated GI_{50}s, other higher or lower concentration ranges of the drug (again made of five, 10-fold dilutions) are also tested. Thus, for each compound tested in the NCI *in vitro* Human Tumor Cell Line Screen, multiple GI_{50} mean graphs (one for each concentration range tested) based on multiple experiments each and with a different content of approximated *versus* calculated (by interpolation) GI_{50}s may exist in the NCI database.

We obtained all of the available GI_{50} mean graph information from the NCI web site for all drugs in our list of compounds with known NSC numbers.[4] For every drug, we recorded the number of concentration ranges tested in the NCI *in vitro* Human Tumor Cell Line Screen, the number of experimental repetitions conducted for each concentration range, and, finally, the number of approximated $Log_{10}GI_{50}$s in each mean graph.

The drug concentration range that produced the mean graph with the smallest number of approximated $Log_{10}GI_{50}$s was used for scoring a drug's activity in the NCI *in vitro* Human Tumor Cell Line Screen, unless a different concentration range existed, with a number of approximated $Log_{10}GI_{50}$s varying <10% from the first but for which more experiments were done.

Preclinical activity in the NCI *in vitro* Human Tumor Cell Line Screen was scored in two different ways: by the mean $Log_{10}GI_{50}$ and by what was termed the activity fraction. For a given drug, in a given tumor type, the mean $Log_{10}GI_{50}$ was computed by averaging the $Log_{10}GI_{50}$s from all of the cell lines of that tumor type in the mean graph corresponding to the most appropriate concentration range. The activity fraction was arbitrarily defined as the number of cell lines of a given tumor type in which the individual $Log_{10}GI_{50}$s were more sensitive to the drug than the average $Log_{10}GI_{50}$ (for all cell lines of all cell types) in the mean graph over the total number of cell lines tested from that tumor type. The activity fraction was also calculated from the mean graph corresponding to the most appropriate concentration range. Overall mean $Log_{10}GI_{50}$s or activity fractions for all four cancer types combined were calculated by averaging the single tumor values.

Table 1 Drugs selected for data collection, NSC numbers are shown, where available

Drug	NSC number
Taxotere	628503
Paclitaxel	125973
Topotecan	609699
Irinotecan	
Rhizoxin	332598
Gemcitabine	
Fazarabine	281272
Teniposide	122819
Menogaril	269148
Fosquidone	D611615
Elsamitrucin	369327
Amonafide	308847
Didemnin B	325319
Suramin	
Raltitrexed	639186
Flavone acetic acid	347512
Epirubicin	256942
CI-921	343499
Trimetrexate	352122
Multitargeted antifol	
Vinorelbine	
Piritrexim	351521
Fotemustine	
CI-980	
Chloroquinoxaline sulfonamide	339004
Ilmofosine	
CI-941	
Tiazofurin	286193
Pyrazine diazohydroxide	361456
Tallimustine	
Crisnatol	

Statistical Analysis

For each preclinical cancer model, 9 Phase II *versus* preclinical activity relationships were examined for a total of 27: relationships by tumor type (disease-oriented approach, 4 relationships/model), predictive ability of one tumor type for the other three tumor types combined (compound-oriented approach, 4 relationships/model), and general predictive ability for all four tumor types combined (1 relationship/model).

Relationships were first examined descriptively with the construction of various Phase II overall activity *versus* preclinical activity scatter plots (Microsoft Excel software). Each point on these scatter plots represented data from one drug for which both Phase II and preclinical activity values had been calculated from literature sources, as described above.

After descriptive evaluation of the data, Spearman rank correlation coefficients were obtained using the SAS software, UNIX version 6.12. A significance test of every correlation coefficient was performed, and the corresponding *P*s were calculated. Spearman rank (nonparametric) correlation coefficients were used because the distributions of the x (preclinical activity) and y (clinical activity) variables were not normal (29).

When multiple comparisons are made within a group of data such as in this work, there is increased possibility that some correlations will come up as statistically significant solely because of chance (false positives). To avoid this, multiple comparison correction methods (*e.g.*, Bonferroni approach) are often used to adjust the significance level to a lower P than conventionally used. However, relying on corrected probabilities increases the possibility that meaningful correlations will be missed (false negatives), making the nature of the scientific work key to the decision to use multiple comparison adjustment methods or not. Because this was an exploratory study, we were willing to accept a higher probability of false positives to ensure that potentially meaningful associations would not be discarded. We therefore did not correct for multiple comparisons and chose a level of significance of 0.05.

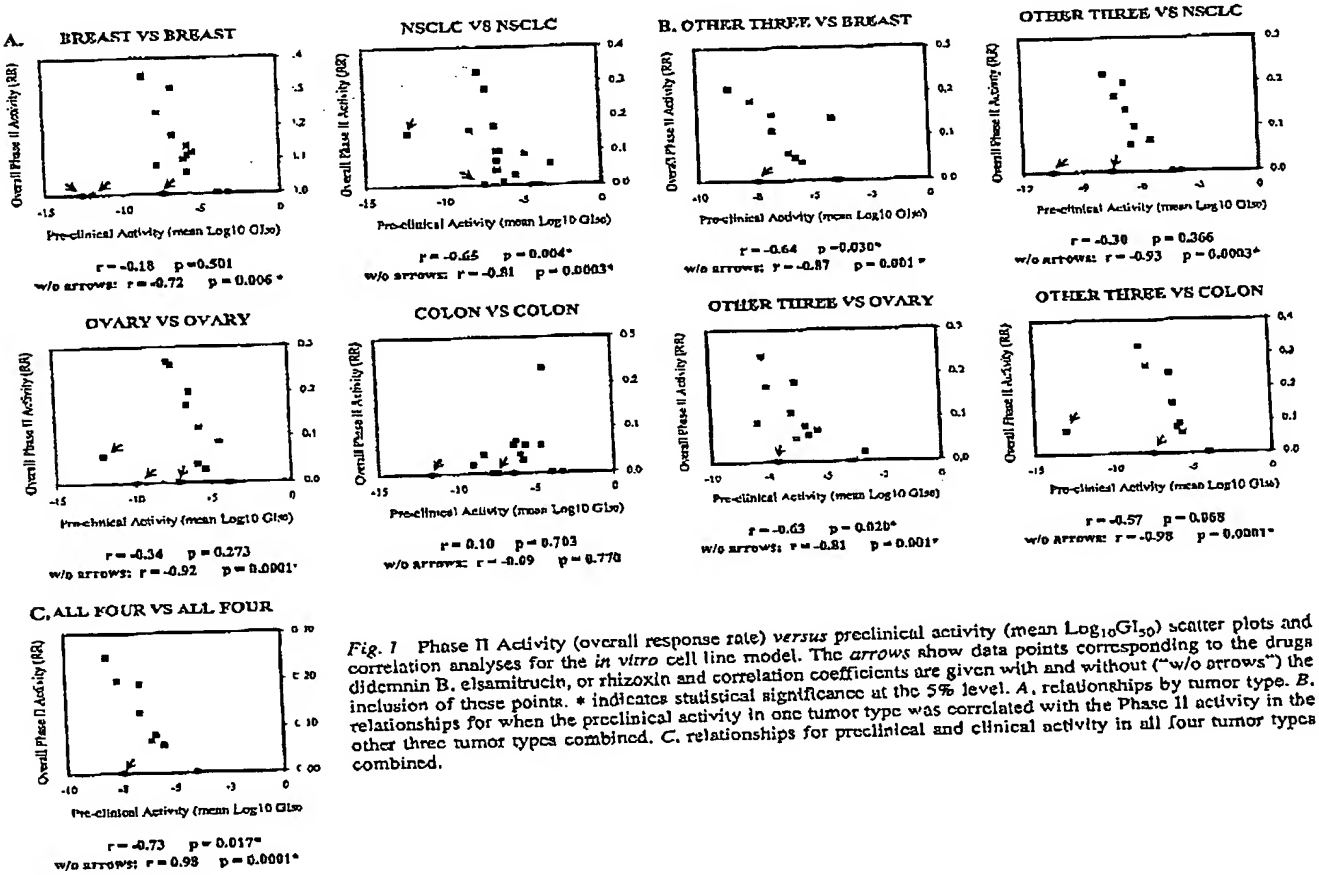


Fig. 1 Phase II Activity (overall response rate) *versus* preclinical activity (mean $Log_{10}GI_{50}$) scatter plots and correlation analyses for the *in vitro* cell line model. The *arrows* show data points corresponding to the drugs didemnin B, elsamitrucin, or rhizoxin and correlation coefficients are given with and without ("w/o arrows") the inclusion of these points. * indicates statistical significance at the 5% level. *A*, relationships by tumor type. *B*, relationships for when the preclinical activity in one tumor type was correlated with the Phase II activity in the other three tumor types combined. *C*, relationships for preclinical and clinical activity in all four tumor types combined.

RESULTS

The Medline and CancerLit databases were searched for cancer drugs (excluding agents with novel targets such as signal transduction or angiogenesis modulators) that had undergone single agent Phase I clinical trial testing either in 1991 or 1992. This search led to 97 drug names. After excluding drugs that were eliminated from additional clinical testing for practical reasons (for example difficulties with the drug formulation), drugs that were specifically developed for a certain type of cancer (as for example hormone-regulating compounds for breast cancer) and drugs that were still the subject of published Phase I studies in 1991 and 1992 despite already being licensed for human use before 1985, a list of 31 agents was obtained (Table 1). After applying the restrictions and criteria mentioned under "Materials and Methods," we extracted from the literature preclinical and Phase II activity information for those agents on four common cancer types, breast, NSCLC, ovary, and colon. Overall, 100 preclinical and 307 Phase II clinical literature references were used spanning the period between 1985 and 2000.

No preclinical data were found for 5 of the 31 drugs researched. Of the 26 drugs remaining, availability of preclinical and Phase II data varied, depending on which preclinical and clinical tumor(s) had been tested and published in each case. Thus, each of the relationships examined had a different number of data points as different subsets of drugs were included. The most data points for any relationship were 17. For six relationships, five or fewer data points were available (relationships with fewer than five data points were not included in the results presented below).

***In Vitro* Cell Line Model.** Fig. 1 shows the Phase II activity *versus* preclinical activity scatter plots and correlation analysis for the *in vitro* cell line model when the mean $Log_{10}GI_{50}$ was used as the measure of preclinical activity. Because the lower the mean $Log_{10}GI_{50}$, the higher the potency of a drug, a negative correlation between mean $Log_{10}GI_{50}$ and Phase II overall RR was expected if the model had a good clinical predictive value. Significant negative correlations were found for NSCLC (Fig. 1*A*), for breast or ovarian cell lines *versus* overall Phase II activity in the other three tumor types

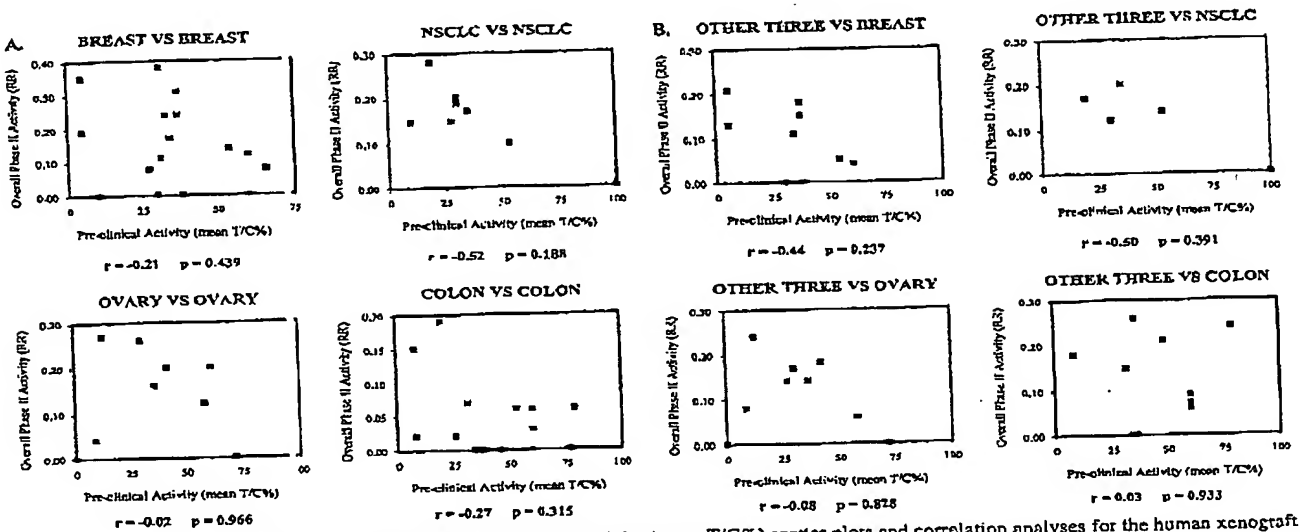


Fig. 2 Phase II activity (overall response rate) *versus* preclinical activity (mean T/C%) scatter plots and correlation analyses for the human xenograft model. *A*, relationships by tumor type. *B*, relationships for when the preclinical activity in one tumor type was correlated with the Phase II activity in the other three tumor types combined.

(Fig. 1*B*), and for preclinical activity *versus* Phase II activity in all four tumor types (Fig. 1*C*).

Although the trends observed with the activity fraction were similar to ones seen for the mean $Log_{10}GI_{50}$ measure, no correlations were statistically significant in this case (data not shown).

Human Xenograft Model. A negative correlation between Phase II RRs and mean T/C% values was expected to be indicative of a good clinical predictive value for the human xenograft model. As shown in Fig. 2, no significant correlations between preclinical and clinical activity were observed for this model in our analysis.

For some of the drugs, preclinical activity calculations were based on multiple human xenografts of the same tumor type (*i.e.*, panels) while for others on only a single xenograft. The relationships in Fig. 2 were reanalyzed, including only the drugs for which preclinical information on more than one human xenograft was available (Fig. 3). The results did not change for breast or colon tumors (compare Fig. 3*A* with Fig. 2*A*). However, the relationship for NSCLC became statistically significant and a highly significant correlation was seen for ovarian cancer (Fig. 3*A*). A near significant correlation was obtained when ovarian human xenograft panels were used to predict clinical activity in the other three tumor types combined (Fig. 3*B*).

Murine Allografts. No significant correlations between preclinical and clinical activity were observed for any of the relationships examined in this study for the murine allograft model (data not shown).

Additional Analyses The scatter plots in Fig. 1 revealed an interesting observation: in every relationship except for colon cancer under the disease oriented approach, an obvious trend toward a negative correlation was evident except for one to three outlier data points (Fig. 1, *arrows*). Interestingly, in all cases, these outlier data points corresponded to the same three drugs, namely elsamitrucin, didemnin B, and rhizoxin.

In an attempt to provide a possible explanation for this observation, we considered the mechanism of action of all drugs that were included in the correlations in Fig. 1. From a total of 18 drugs (Table 2), 5, namely, elsamitrucin, didemnin B, rhizoxin, flavone acetic acid, and fosquidone, were distinct in that they seemed to act through mostly unknown pathways that were not the typical DNA-based mechanisms of action of cytotoxic cancer agents. Thus, although flavone acetic acid and fosquidone fitted the rest of the data, there seemed to be a plausible mechanistic basis for the outlier behavior of the data points for elsamitrucin, didemnin B, and rhizoxin. In fact, exclusion of these three drugs led to highly significant correlations in all cases except for the same tumor relationship in colon cancer (Fig. 1, correlation coefficients and *P*s for "w/o arrows"). It should be noted that none of the relationships examined for the human xenograft models (Figs. 2 and 3) included elsamitrucin, didemnin B, or rhizoxin as data points.

Because of the intriguing results obtained with the human NSCLC and ovarian xenograft panels in Fig. 3*A*, a more detailed examination of these panels was pertained. As seen in Figs. 4*A* and 5*A*, the 6 ovarian and 7 NSCLC xenograft panels differed both in the numbers (minimum of 6 and maximum of 13 for ovary and minimum of 2 and maximum of 8 for NSCLC) and the identity of the xenografts that they contained. Analysis by grade/histology was hindered by lack of complete information on all xenografts. However, some patterns appeared distinguish-

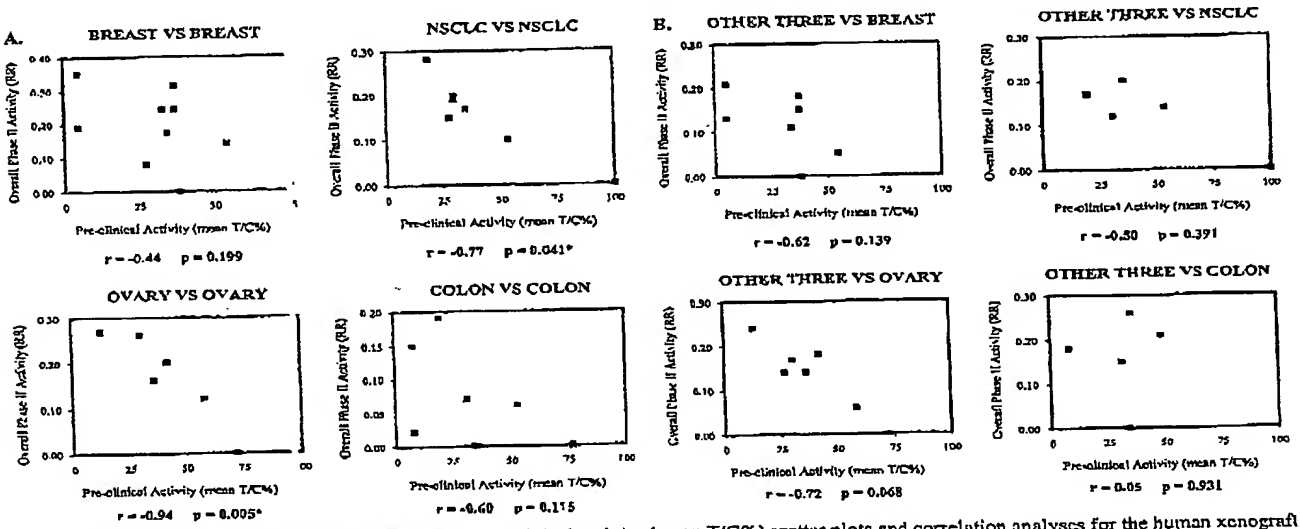


Fig. 3 Phase II activity (overall response rate) *versus* preclinical activity (mean T/C%) scatter plots and correlation analyses for the human xenograft model. Only data points for which two or more human xenografts were used to generate the preclinical activity values are shown. * indicates statistical significance at the 5% level. *A*, relationships by tumor type. *B*, relationships for when the preclinical activity in one tumor type was correlated with the Phase II activity in the other three tumor types combined.

Table 2 Mechanisms of action of drugs used in clinical vs. pre-clinical correlations for the *in vitro* cell line model (Fig. 1)
Atypical cytotoxics are shown in bold.

Drug	Mechanism of action
Amonafide	DNA intercalator
CI-921	Acts on topoisomerase II
Didemnin B	Not understood. Believed to act on protein synthesis
Elsamitrucin	Not understood. It has been observed to inhibit topoisomerase I and II in *in vitro* experiments (relevance to *in vivo* uncertain). In cells in culture it has been observed to cause a cytostatic effect.
Epirubicin	Attaches to DNA at G bases
Fazarabine	Probably inhibits DNA synthesis by incorporation into DNA.
Flavone acetic acid	Has antivascular action in mice (probably not applicable to humans). Also believed to induce cell cycle arrest by generating reactive oxygen species that act on DNA.
Menogaril	Causes cleavage of double-stranded DNA by inhibiting topoisomerase II
Piritrexim	Inhibits dihydrofolate reductase
Rhizoxin	Not fully understood. May interact with tubulin (different binding site than taxoids) and lead to cell cycle arrest. Also observed to act as an angiogenesis inhibitor.
Taxol	Microtubule destabilizing agent that causes apoptosis
Taxotere	Microtubule destabilizing agent that causes apoptosis
Teniposide	DNA synthesis inhibition by stabilization of cleavable DNA complexes
Topotecan	Topoisomerase I inhibitor
Trimetrexate	Antifolate
Fosquidone	Unknown. Not a DNA binder or a topoisomerase inhibitor
Tomudex	Thymidylate synthase inhibitor
Tiazofurin	Inhibits 5'-phosphodehydrogenase, the rate-limiting enzyme for guanine ribonucleotide synthesis

able. All ovarian panels contained 10–20% undifferentiated tumors and also included both poorly differentiated and moderately differentiated subtypes (Fig. 4*B*). For NSCLC, all panels included adenocarcinoma xenografts with a frequency of >30% (Fig. 5*B*). These observations suggested that the frequency of histological/grade subtypes within a xenograft panel may be an important determinant of clinical predictivity rather than the number or the nature of the xenografts.

In an attempt to explore this hypothesis and to further examine the validity of the results obtained for ovarian cancer and NSCLC in Fig. 3*A*, the literature was reviewed for additional data. Six more agents with known overall Phase II RRs in

A.

NAME	HISTOLOGY / GRADE	DATA POINTS (DRUGS) EPIRUBICIN	POSQUIDONE	GEMCITABINE	MENOGARIL	TAXOTERE	PACLITAXEL
MRI-H-207	undifferentiated	•	•	•	•	•	•
A2780	undifferentiated	•		•	•		
Ov.He	mod. diff., mucinous	•	•		•		
Ov.Me	carcinosarcoma	•	•	•	•		
OvRiC	mod. diff., serous	•	•		•	•	
Fma	poorly diff., mucinous	•	•	•	•	•	
Ov.Pe	mod. diff., mucinous	•	•		•		
Fco	clear cell sarcoma	•					
T17	cystoadenocarcinoma	•					
T385	adenocarcinoma	•					
OvGR	mod. diff., mucinous		•			•	
Fko	mod. diff., serous		•	•			
OvG1	poorly diff., serous		•			•	•
OVCAR-3	adenocarcinoma			•		•	•
A121a	?					•	•
HOC18	poorly diff., serous					•	•
HOC22	poorly diff., serous					•	•
A2780/DDP	undifferentiated					•	•
A2780/DX	undifferentiated					•	•
SKOV-3	adenocarcinoma					•	•
1° ovary 1	cystoadenocarcinoma					•	•
1° ovary 2	dediff. serous adenoc.					•	•
IGROV 1	moderately diff.					•	•
OVCAR-8	poorly diff. adenoc.					•	•
OVCAR-5	adenocarcinoma					•	•
OvSh	poorly diff., serous					•	•
HOC22-S	poorly diff., serous					•	•
TOTAL NO.		10	10	6	8	10	13

B.

HISTOLOGY/GRADE	FREQUENCIES IN HUMAN OVARIAN XENOGRAFT PANELS					
HISTOLOGY / GRADE	EPIRUBICIN NO. (%)	POSQUIDONE NO. (%)	GEMCITABINE NO. (%)	MENOGARIL NO. (%)	TAXOTERE NO. (%)	PACLITAXEL NO. (%)
undifferentiated	2 (20)	1 (10)	1 (17)	2 (25)	1 (10)	3 (23)
mod. diff., mucinous	2 (20)	3 (30)	2 (33)	2 (25)	1 (10)	0 (0)
mod. diff., serous	1 (10)	2 (20)	2 (33)	1 (12.5)	1 (10)	0 (0)
poorly diff., mucinous	1 (10)	1 (10)	0 (0)	1 (12.5)	1 (10)	0 (0)
poorly diff., serous	0 (0)	1 (10)	0 (0)	0 (0)	4 (40)	2 (15)
unspecified	4 (40)	2 (20)	1 (17)	2 (25)	2 (20)	8 (62)
TOTAL	10 (100)	10 (100)	6 (100)	8 (100)	10 (100)	13 (100)

Fig. 4 Human ovarian xenograft panels for the six data points (drugs) used in the "Ovary *versus* Ovary" relationship in Fig. 3*A*. *A*, names and histology/grade (? = unknown, mod. diff. = moderately differentiated, poorly diff. = poorly differentiated, dediff. = dedifferentiated, adnerc = adenocarcinoma) of all of the xenografts tested. Inclusion of a particular xenograft in one of the panels is shown by a "+" sign in the corresponding row and under the appropriate drug column. *B*, histology/grade subtypes in the human ovarian xenograft panels by number and percentage.

previously treated patients with ovarian cancer were found. Five and one of these compounds had been tested in a panel of 15 and 6 human ovarian xenografts, respectively (26, 30), which fitted the histology/grade patterns identified in Fig. 4*B*. Fig. 6*A* lists the names and Phase II RRs (31–56) of these additional drugs together with the six compounds that were included in the analysis in Fig. 3*A*. Fig. 6, *A* and *B*, also shows mean T/C% values scatter plots and statistical analyses for two cases: first, for when all of the available xenograft information was used, and second, for when mean T/C% calculations were based, where possible, on the arithmetically smallest panel, namely the one used for gemcitabine in Fig. 4. Highly significant correlations were obtained in both cases (Fig. 6*B*).

For NSCLC information on two additional agents was found: amsacrine [mean T/C% of 62 (26) and Phase II RR equal to 0.06 (31)] and doxorubicin [mean T/C% of 47 (26) and Phase II RR equal to 0.12 (32)]. Both had been tested in NSCLC human xenograft panels that included all three histological subtypes and had adenocarcinoma contents of 29 and 33%, respectively. As for ovarian cancer, those two additional data points (Fig. 5*C*, *arrows*) enhanced the statistical significance of the relationship observed in Fig. 3*A*.

DISCUSSION

A literature-based, retrospective study was conducted to examine the clinical predictive value of three widely used preclinical cancer models, namely, the *in vitro* human tumor cell line, the human xenograft, and the murine allograft models. Four solid tumor types were selected, breast, NSCLC, ovary and colon, and data on a set of 31 anticancer agents (excluding agents with novel targets such as signal transduction or angiogenesis modulators) were collected. Preclinical activity in each model was correlated with RRs in Phase II clinical trials by tumor type (disease-oriented approach) in the case when one preclinical tumor type was used as a predictor of overall clinical activity in the other three tumor types combined (compound-oriented approach) and for all four tumor types together.

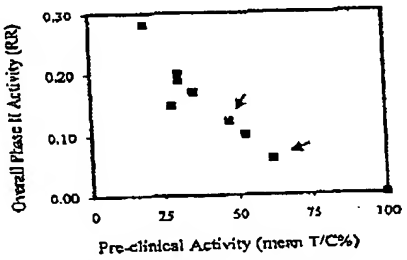
A.

XEN. NAME	XENOGRAFT HISTOLOGY	DRUGS						
		EPI	FAZ	GEM	IRINO	PACLIT	TOPO	VINRLB
T222	squamous cell	+						
T291	adenocarcinoma	+						
UCLA-P3	adenocarcinoma		+					
ACC01-U-78	squamous cell		+					
NCI-H460	large cell			+	+	+	+	+
A549	adenocarcinoma			+	+	+	+	+
CaLu-6	adenocarcinoma			+				
H-74	?			+				
LC-376	?			+				+
QG-56	squamous cell				+	+		
NCI-H23	adenocarcinoma				+	+	+	
NCI-H226	squamous cell				+	+		
MV-522	adenocarcinoma					+		
CaLu-3	adenocarcinoma					+		
1° NSCLC	adenocarcinoma					+		
L2987	adenocarcinoma							+
L-27	adenocarcinoma							+
LC-06	large cell							+
LU-65	large cell							+
PC-12	adenocarcinoma							+
LU-99	large cell							+
TOTAL NO.		2	2	5	5	8	3	8

B.

HISTOLOGY FREQUENCY IN HUMAN NSCLC XENOGRAFT PANELS						
HISTOLOGY	EPI NO. (%)	FAZ NO. (%)	GEM NO. (%)	IRINO NO. (%)	PACLIT NO. (%)	TOPO NO. (%)	VINORLB NO. (%)
adenocarcinoma	1 (50)	1 (50)	2 (40)	2 (40)	6 (75)	1 (33.3)	3 (37.5)
large cell	0 (0)	0 (0)	1 (20)	1 (20)	1 (12.5)	1 (33.3)	4 (50)
squamous cell	1 (50)	1 (50)	0 (0)	2 (40)	1 (12.5)	1 (33.3)	1 (12.5)
unknown			2 (40)				
TOTAL	2 (100)	2 (100)	5 (100)	5 (100)	8 (100)	3 (100)	8 (100)

C. NSCLC VS NSCLC (ADDITIONAL DATA)



r = -0.90 p = 0.001*

Fig. 5 Human NSCLC xenograft panels for the seven data points (drugs) used in the NSCLC *versus* NSCLC relationship in Fig. 3A. *A*, drug names (EPI = epirubicin, FAZ = fazarabine, GEM = gemcitabine, IRINO = irinotecan, PACLIT = paclitaxel, TOPO = topotecan, VINRLB = vinorelbine) and histological subtype: (? = unknown) of all of the xenografts tested. Inclusion of a particular xenograft in one of the panels is shown by a "+" sign in the corresponding row and under the appropriate drug column. *B*, histological subtypes in the human NSCLC xenograft panels by number and percentage. *C*, scatter plot and correlation analysis for the same tumor clinical *versus* preclinical activity relationship in NSCLC, including the seven drugs in Fig. 6A as well as two additional agents, doxorubicin and amsacrine (data points shown with *arrows*), with known NSCLC Phase II and human xenograft activities

Colon cancer was the only site for which a disproportional amount of clinically active *versus* inactive agents were identified: only 3 drugs with Phase II RRs > 0.15 and 8 with ≤0.10 (Figs. 1–3). However, this was likely a reflection of the lack of clinically effective drugs for this tumor type rather than the result of selection and publication bias.

When the mean Log $_{10}GI_{50}$ measure of preclinical activity was used, the *in vitro* cell line model was found to be predictive of Phase II clinical performance for NSCLC under the disease-oriented approach in breast and ovarian cancers under the compound-oriented approach and in the case of all four tumor types together. Highly significant correlations were observed in all cases, except colon cancer, when three consistent outlier data points corresponding to the mechanistically nontypical cytotoxic agents didemnin B, elsamitrucin, and rhizoxin were excluded in exploratory analysis. Thus, the *in vitro* cell line model

A.

	DRUG	PHASE II RESPONSE RATE	HUMAN OVARIAN XENOGRAFT MEAN T/C%	
			ALL TESTED	GEMCITABINE PANEL
STUDY DRUGS	EPIRUBICIN	0.20	42	-
	FOSQUIDONE	0.00	72	-
	GEMCITABINE	0.16	36	36
	MENOGARIL	0.12	58	-
	PACLITAXEL	0.26	30	-
	TAXOTERE	0.27	13	-
ADDITIONAL DRUGS	DOXORUBICIN	0.19 [366]	47 [30]	47 [30]
	AMSACRINE	0.05 [9]	75 [30]	-
	CISPLATIN	0.25 [7,30]	41 [30]	46 [30]
	HEXAMETHYL-MELAMINE	0.19 [9,44]	28 [30]	31 [30]
	METHOTREXATE	0.09 [9,44]	76 [30]	84 [30]
	5-FU	0.10 [9,44]	71 [30]	71 [30]

B.

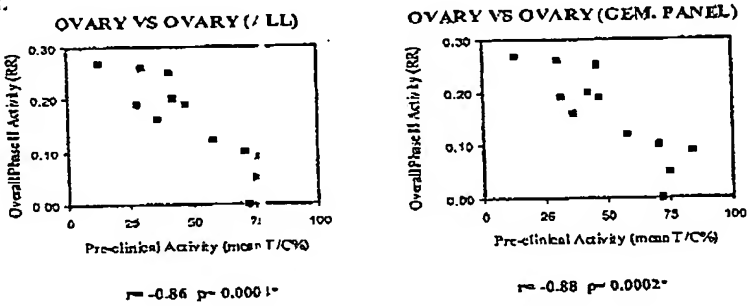


Fig. 6 *A*, preclinical and Phase II clinical activity data for ovarian cancer, including the six drugs in Fig. 3*A* ("Study Drugs") as well as an additional six drugs ("Additional Drugs") with known ovarian Phase II and human xenograft activities. Literature references are shown in superscript font. *B*, scatter plots and correlation analysis for the same tumor clinical *versus* preclinical activity relationship in ovarian cancer based on the data in Fig. 5*A*. Analysis was done for (*a*) when all of the xenografts were included in preclinical activity calculations ("All") and (*b*) when only the six xenografts in the gemcitabine panel were used for preclinical activity calculations, where possible ("Gem. Panel"). Stars indicate statistical significance at the 5% level.

might be predictive in the case of typical cytotoxic cancer agents but might fail to provide reliable information for at least some of the noncytotoxic cancer drugs. Additional studies are needed to explore this observation.

The fact that drug potency (mean $Log_{10}GI_{50}$), a pharmacological measure, was found to be predictive of Phase II performance was somewhat surprising but has been noted previously; a recent study by Johnson *et al.* (18) demonstrated a highly significant correlation between potency in the NCI human tumor cell line screen and activity in the hollow fiber assay. Pharmacological considerations (pharmacological differences between the species) might provide a possible explanation why some anticancer agents appear effective in *in vivo* mouse models but fail to show efficacy in Phase II trials. Experience with some agents (57) has shown that the maximum-tolerated dose in mouse can be higher than in humans, presumably because of an intrinsic ability of mouse cells to tolerate higher drug doses and/or more efficient elimination in the mouse.

In contrast to the *in vitro* cell line, our results suggest that the murine allograft model, as used in this analysis, is not predictive of clinical Phase II performance. This is in agreement with the conclusions from a large body of information originating from the NCI screening programs in use from 1975 to 1990 (5–8, 10–12).

The human xenograft model showed good tumor-specific predictive value for NSCLC and ovarian cancers when panels of xenografts were used. However, it failed to adequately predict clinical performance both in the disease and compound-oriented settings for breast and colon tumors. The results with breast cancer were in agreement with a recent study (18) but were contradictory to the work reported by Bailey *et al.* (20). Inoue *et al.* (21), and Mattern *et al.* (24). However, given that the latter studies did not use formal statistical methods, our conclusions may be more robust. The results for ovarian cancer were in agreement with studies by Taetle *et al.* (23) and Mattern *et al.* (24) but contradicted the conclusions of the recent NCI United States study by Johnson *et al.* (18). Our results for NSCLC were consistent with the observations from all previous studies that examined same tumor correlations in this cancer type (18, 24).

For NSCLC and ovarian cancer patients, a panel of xenografts was more predictive than single xenografts confirming preliminary observations by Bellet *et al.* (19).

In an effort to identify the properties that may render an ovarian or NSCLC human xenograft panel predictive of Phase II drug performance, common characteristics were sought. There was no similarity in number and only limited overlap in identity of xenografts between same tumor type panels. However, certain patterns in histology/grade content were found. These ob-

servations suggest that the relative histology/grade content rather than the number or identity of xenografts within a panel may be the important determinant of clinical predictivity. To our knowledge, no other study has attempted to identify ovarian or NSCLC human xenograft panel features that might lead to accurate predictions of a drug's Phase II performance.

This is the only study that has examined the clinical predictive value of three preclinical cancer models together and thus allows for direct comparisons between them. The results suggest that the human xenograft model is more predictive than its murine allograft counterpart and that the *in vitro* cell line model is of, at least, equivalent usefulness to the human xenograft model.

The NCI work with cancer drug screening programs from 1955 to 1990 (Refs. 5–8, 10–12; leukemia-based preclinical, compound-oriented screens preferentially yielding compounds active against hematological malignancies) in combination with our work and recent conclusions by Johnson *et al.* (Ref. 18; statistically significant results under the compound-oriented approach for some solid tumors) suggest that the compound-oriented strategy may be successful when used only within solid tumors or only within hematological malignancies but not when the two disease groups are considered together.

In general, our results suggest that the *in vitro* human tumor cell line and the human xenograft models might have good clinical predictive value in some solid tumors (such as ovary and NSCLC) under both the disease and compound-oriented strategies, as long as an appropriate panel of tumors is used in preclinical testing.

In conclusion, given the results in this study and those of others (6, 7, 10–12), continued use of the murine allograft model in drug development may not be justified. The work presented here argues for emphasis to be placed on *in vitro* cell lines (in the context of the NCI Human Tumor Cell Line Screen) and appropriate panels of the human xenograft model.

Recent years have seen an explosion in the molecular understanding of cancer, which has led to the development of not only more effective cytotoxic cancer drugs but of potentially cytostatic or antimetastatic agents as well. The future preclinical and clinical development of traditional cytotoxic compounds will likely follow similar procedures with those practiced today, and in that sense, the present findings could contribute to the more efficient discovery of such agents. However, the existing cancer models and parameters of activity in both the preclinical and clinical settings may have to be redesigned to fit the mode of action of the novel cytostatic, antimetastatic, antiangiogenesis, or immune response-modulating agents (58). In the preclinical cancer model front, the case is being made for the use of the orthotopic mouse xenograft and transgenic models (59–61) because those are thought to more accurately simulate human disease, especially in terms of growth characteristics and metastatic behavior. New end points of preclinical activity are contemplated such as the demonstration that a new molecule truly hits the intended molecular target (58). In Phase II clinical trials, there is a growing effort toward validating new surrogate endpoints of drug efficacy (58). The next decade will probably answer many of the questions regarding the effectiveness of these novel agents and will likely define a new role for traditional cytotoxic therapies, but it will also bring new challenges in terms of preclinical predictors of activity.

REFERENCES

1. Curt, G. A. The use of animal models in cancer drug discovery and development. Stem Cells, *12:* 23–29, 1994.
2. Goldin, A., Serpick, A. A., and Mantel, N. Experimental screening procedures and clinical predictability value. Cancer Chemother. Rep., *50:* 173–218, 1966.
3. Venditti, J. M. Drug evaluation branch program: report to screening contractors. Cancer Chemother. Rep., *5:* 1–4, 1975.
4. Venditti, J. M. Preclinical drug development: rationale and methods. Semin. Oncol., *8:* 349–353, 1981.
5. Venditti, J. M., Wesley, R. A., and Plowman, J. Current NCI preclinical antitumor screening *in vivo*: results of tumor panel screening, 1976–1982 and future directions. Adv. Pharmacol. Chemother., *20:* 1–19, 1984.
6. Venditti, J. M. The National Cancer Institute antitumor drug discovery program, current and future perspectives: a commentary. Cancer Treat. Rep., *67:* 767–772, 1983.
7. Staquet, M. J., Byar, D. P., Green, S. B., and Rozencweig, M. Clinical predictivity of transplantable tumor systems in the selection of new drugs for solid tumors: rationale for a three-stage strategy. Cancer Treat. Rep., *67:* 753–765, 1983.
8. Goldin, A., Venditti, J. M., MacDonald, J. S., Muggia, F. M., Henney, J. E., and De Vita V. T. Current results of the screening program at the division of cancer treatment, National Cancer Institute. Eur. J. Cancer, *17:* 129–142, 1981.
9. Marsoni, S., and Wittes, R. Clinical development of anticancer agents: A National Cancer Institute perspective. Cancer Treat. Rep., *68:* 77–85, 1983.
10. De Vita, V. T., and Schein, P. S. The use of drugs in combination for the treatment of cancer N. Engl. J. Med., *288:* 998–1006, 1973.
11. Zubrod, C. G. Chemical control of cancer. Proc. Natl. Acad. Sci. USA, *69:* 1042–1047, 1972.
12. Marsoni, S., Hoth, D., Simon, R., Leyland-Jones, B., De Rosa, M., and Wittes, R. E. Clinical drug development: an analysis of Phase II trials, 1970–1985. Cancer Treat. Rep., *71:* 71–80, 1987.
13. Shoemaker, R. H., Monks, A., Alley, M. C., Scudiero, D. A., Fine, D. L., McLemore, T. L., Abbott, B. J., Paull, K. D., Mayo, J. G., and Boyd, M. R. Development of human tumor cell line panels for use in disease-oriented drug screening. Prog. Clin. Biol. Res., *276:* 265–286, 1988.
14. Skehan, P., Storeng, R., Scudiero, D., Monks A., McMahon, J., Vistica, D., Warren, J. T., Bokesch, H., Kenney, S., and Boyd, M. R. New colorimetric cytotoxicity assay for anticancer drug screening. J. Natl. Cancer Inst. (Bethesda), *82:* 1107–1112, 1990.
15. Rubinstein, L. V., Shoemaker, R. H., Paull, K. D., Simon, R. M., Tosini, S., Skehan, P., Scudiero, D., Monks, A., and Boyd, M. R. Comparison of *in vitro* anticancer drug screening data generated with a tetrazolium assay *versus* a protein assay against a diverse panel of human tumor cell lines. J. Natl. Cancer Inst. (Bethesda), *82:* 1113–1118, 1990.
16. Alley, M. C., Scudiero, D., Monks, A., Hursey, M. L., Czerwinski, M. J., Fine, D. L., Abbott, B. J., Mayo, J. G., Shoemaker, R. H., and Boyd, M. R. Feasibility of drug screening with panels of human tumor Cell lines using a microculture tetrazolium assay. Cancer Res., *48:* 589–601, 1988.
17. Monks, A., Scudiero, D., Skehan, P., Shoemaker, R. H., Paull, K. D., Vistica, D., Hose, C., Langley, J., Cronise, P., Vaigro-Wolff, A., Gray-Goodrich, M., Campbell, H., Mayo, J., and Boyd, M. R. Feasibility of a high-flux anticancer drug screen using a diverse panel of cultured human tumor cell lines. J. Natl. Cancer Inst. (Bethesda), *83:* 757–766, 1990.
18. Johnson, J. I., Decker, S., Zaharevitz, D., Rubinstein, L. V., Venditti, J. M., Schepartz, S., Kalyandrug, S., Christin, M., Arbuck, S., Hollingshead, M., and Sausville, E. A. Relationships between drug

activity in NCI preclinical *in vitro* and *in vivo* models and early clinical trials. Br. J. Cancer, *84:* 1424–1431, 2001.

19. Bellet, R. E., Danna, V., Mastrangelo, M. J., and Berd, D. Evaluation of a "nude" mouse-human tumor panel as a predictive secondary screen for cancer chemotherapy agents. J. Natl. Cancer Inst. (Bethesda), *63:* 1185–1187, 1979.

20. Bailey, M. J., Gazet, J-C., Smith, I. E., and Steel, G. G. Chemotherapy of human breast carcinoma xenografts. Br. J. Cancer, *42:* 530–536, 1980.

21. Inoue, K., Fujimoto, S., and Ogawa, M. Antitumor efficacy of seventeen anticancer drugs in human breast cancer xenograft (MX-1) transplanted in nude mice. Cancer Chemother. Pharmacol., *10:* 182–186, 1983.

22. Steel, G. G., Courtenay, V. D., and Peckham M. J. The response to chemotherapy of a variety of human tumor xenografts. Br. J. Cancer, *47:* 1–13, 1983.

23. Tuttle, R., Rosen, F., Abramson, I., Venditti, J., and Howell, S. Use of nude mouse xenografts as preclinical drug screens: *in vivo* activity of established chemotherapeutic agents against melanoma and ovarian carcinoma xenografts. Cancer Treat. Rep., *71:* 297–304, 1987.

24. Mattern, J., Bak, M., Hahn, E. W., and Volm, M. Human tumor xenografts as models for drug testing. Cancer Metastasis Rev., *7:* 263–284, 1988.

25. Boven, E., Winograd, B., Fodstad, O., Lobbezzoo, M. W., and Pinedo, H. M. Preclinical Phase II studies in human tumor lines: a European multicenter study. Eur. J. Cancer, *24:* 567–573, 1988.

26. Boven, E., Winograd, B., Berger, D. P., Dumant, M. P., Braakhuis, B. J. M., Fodstad, O., Langdon, S., and Piebig, H. H. Phase II preclinical drug screening in human tumor xenografts: a first European multicenter collaborative study. Cancer Res. *52:* 5940–5947, 1992.

27. Langdon, S., Hendriks, H. R., Braakhuis, B. J. M., Pratesi, G., Berger, D. P., Fodstad, O., Piebig, H., and Boven, E. Preclinical Phase II studies in human tumor xenografts: a European multicentet follow-up study. Ann. Oncol., *5:* 415–422, 1994.

28. Paull, K. D., Shoemaker, R. H., Hodes, L., Monks, A., Scudiero, D., Rubinstein, L. V., Plowman, J., and Boyd, M. R. Display and analysis of patterns of differential activities of drugs against human tumor cell lines: development of mean graph and COMPARE algorithm. J. Natl. Cancer Inst. (Bethesda), *81:* 1088–1092, 1989.

29. Rosner, B. Non-parametric Methods. *In:* Fundamentals of Biostatistics. 4th ed., p. 575. Belmont CA: Duxbury Press, Wadsworth Publishing Company, 1994.

30. Kolfscoten, G. M., Pinedo, H. M., Scheffer, P. G., Schluper, H. M. M., Erkelens, C. A. M., and Boven, E. Development of a panel of 15 human ovarian cancer xenografts for drug screening and determination of the role of the glutathione detoxification system. Gynecol. Oncol., *76:* 362–368, 2000.

31. Loute, A. C., and Issell, B. F. Amsacrine (AMSA): a clinical review. J. Clin. Oncol., *3:* 562–592, 1985.

32. Minna, J. D., Pass, H., Glatstein, E., and Inde, D. Cancer of the lung. *In:* V. T. DeVita, S. Hellman, and S. A. Rosenberg (eds.), Cancer Principles and Practice of Oncology, pp. 660 (Table 22–48). Philadelphia: J. B. Lippincott, 1989.

33. Gordon, A. N., Granai, C. O., Rose, P. G., Hainsworth, J., Lopez, A., Weissman, C., Rosales, R., and Sharpington, T. Phase II study of liposomal doxorubicin in platinum- and paclitaxel-refractory epithelial ovarian cancer. Clin. Oncol., *8:* 3093–3100, 2000.

34. Israel, V. P., Garcia, A. A., Roman, L., Muderspach, L., Burnett, A., Jeffers, S., and Muggia, F. M. Phase II study of liposomal doxorubicin in advanced gynecologic cancers. Gynecol. Oncol., *78:* 143–147, 2000.

35. Muggia, F. M., Hainsworth, J. D., Jeffers, S., Miller, P., Groshen, S., Tan, M., Roman, L., Uziely, B., Muderspach, L., Garcia, A., Burnett, A., Greco, F. A., Morrow, C. P., Paradiso, L. J., and Liang, L. J. Phase II study of liposomal doxorubicin in refractory ovarian cancer: antitumor activity and toxicity modification by liposomal encapsulation. J. Clin. Oncol., *15:* 987–993, 1997.

36. Blum, R. H., and Carter, S. K. Adriamycin: a new anticancer drug with significant clinical activity. Ann. Intern. Med., *80:* 249–259, 1974.

37. Thigpen, J. T., Lagasse, L., Homesley, H., and Blessing, J. A. *Cis*-platinum in the treatment of advanced or recurrent adenocarcinoma of the ovary. A Phase II study of the Gynecology Oncology. Group. Am. J. Clin. Oncol., *6:* 431–435, 1983.

38. Wiltshaw, E., and Kroner, T. Phase II study of *cis*-dichlorodiammineplatinum(II) (NSC-119875) in advanced adenocarcinoma of the ovary. Cancer Treat. Rep., *60:* 55–60, 1976.

39. Bonomi, P. D., Mladineo, J., Morrin, B., Wilbanks, G., Jr., and Slayton, R. E. Phase II trial of hexamethylmelamine in ovarian carcinoma resistant to alkylating agents. Cancer Treat. Rep., *63:* 137–138, 1979.

40. Johnson, B. L., Fisher, R. I., Bender, R. A., DeVita, V. T., Jr., Chabner, B. A., and Young, R. C. Hexamethylmelamine in alkylating agent-resistant ovarian carcinoma. Cancer (Phila.), *42:* 2157–2161, 1978.

41. Omura, G. A., Blessing, J. A., Morrow, C. P., Buchsbaum, H. J., and Homesley, H. D. Follow-up on a randomized trial of melphalan (M) *vs.* melphalan plus hexamethylamine (M+H) *versus* Adriamycin plus cyclophosphamide (A+C) in advanced ovarian carcinoma. Proc. Am. Assoc. Cancer Res., *22:* 470, 1981.

42. Bolis, G., D'Incalci, M., Belloni, C., and Mangioni, C. Hexamethylmelamine in ovarian cancer resistant to cyclophosphamide and Adriamycin. Cancer Treat. Rep., *63:* 1375–1377, 1979.

43. Manetta, A., MacNeill, C., Lyter, J. A., Scheffler, B., Podczaski, E. S., Larson, J. E., and Schein, P. Hexamethylmelamine as a single second-line agent in ovarian cancer. Gynecol. Oncol., *36:* 93–96, 1990.

44. Rosen, G. F., Lurain, J. R., and Newton, M. Hexamethylmelamine in ovarian cancer after failure of cisplatin-based multiple-agent chemotherapy. Gynecol. Oncol., *27:* 173–179, 1987.

45. Parker, L. M., Griffiths, C. T., Yankee, R. A., Knapp, R. C., and Canellos, G. P. High-dose methotrexate with leucovorin rescue in ovarian cancer: a Phase II study. Cancer Treat. Rep., *63:* 275–279, 1979.

46. Barlow, J. J., and Piver, M. S. Methotrexate (NSC-740) with citrovorum factor (NSC-3590) rescue, alone and in combination with cyclophosphamide (NSC-26271), in ovarian cancer. Cancer Treat. Rep., *60:* 527–533, 1976.

47. Morgan, R. J., Jr., Speyer, J., Doroshow, J. H., Margolin, K., Raschko, J., Sorich, J., Akman, S., Leong, L., Somlo, G., and Vasilev, S. Modulation of 5-fluorouracil with high-dose leucovorin calcium: activity in ovarian cancer and correlation with CA-125 levels. Gynecol. Oncol., *58:* 79–85, 1995.

48. Markman, M., Reichman, B., Hakes, T., Hoskins, W., Rubin, S., Jones, W., and Lewis, J. L., Jr. Intraperitoneal chemotherapy as treatment for ovarian carcinoma and gastrointestinal malignancies: the Memorial Sloan-Kettering Cancer Center experience). Acta Med. Austriaca, *16:* 65–67, 1989.

49. Prefontaine, M., Donovan, J. T., Powell, J. L., and Buley, L. Treatment of refractory ovarian cancer with 5-fluorouracil and leucovorin. Gynecol. Oncol., *61:* 249–252, 1996.

50. Kamphuis, J. T., Huider, M. C., Ras, G. J., Verhagen, C. A., Kateman, I., Vreeswijk, J. H., and Burghouts, J. T. High-dose 5-fluorouracil and leucovorin as second-line chemotherapy in patients with platinum-resistant epithelial ovarian cancer. Cancer Chemother. Pharmacol., *37:* 190–192, 1995.

51. Burnett, A. F., Barter, J. P., Potkul, R. K., Jarvis, T., and Barnes, W. A. Ineffectiveness of continuous 5-fluorouracil as salvage therapy for ovarian cancer. Am. J. Clin. Oncol., *17:* 490–493, 1994.

52. Reed, E., Jacob, J., Ozols, R. F., Young, R. C., and Allegra, C. 5-Fluorouracil (5-FU) and leucovorin in platinum-refractory advanced stage ovarian carcinoma. Gynecol. Oncol., *46:* 326–329, 1992.

53. Ozols, R. F., Speyer, J. L., Jenkins, J., and Myers, C. E. Phase II trial of 5-FU administered ip to patients with refractory ovarian cancer. Cancer Treat. Rep., *68:* 1229–1232, 1984.

54. Long, H. J., III, Nelimark, R. A., Su, J. Q., Garneau, S. C., Levitt, R., Goldberg, R. M., Poon, M. A., and Kugler, J. W. Phase II evaluation of 5-fluorouracil and low-dose leucovorin in cisplatin-refractory advanced ovarian carcinoma. Gynecol. Oncol., *54:* 180–183, 1994.

55. Look, K. Y., Muss, H. B., Blessing, J. A., and Morris, M. A. Phase II trial of 5-fluorouracil and high-dose leucovorin in recurrent epithelial ovarian carcinoma. A Gynecology Oncology Group Study. Am. J. Clin. Oncol., *18:* 19–22, 1995.

56. De Graeff, A., van Hoef, M. E., Tjin, P., Heintz, A. P., and Neijt, J. P. Continuous infusion of 5-fluorouracil in ovarian cancer patients refractory to cisplatin and carboplatin. Ann. Oncol., *2:* 691–692, 1991.

57. Volpe, D. A., Tomaszewski, J. E., Parchment, R. E., Garg, A., Flora, K. P., Murphy, M. J., and Grieshaber, C. K. Myelotoxic effects of the bifunctional alkylating agent bizelesin in human, canine and murine myeloid progenitor cells. Cancer Chemother. Pharmacol., *39:* 143, 1996.

58. Gelmon, K. A., Eisenhauer, E. A., Harris, A. L., Ratain, M. J., and Workman, P. Anticancer agents targeting signaling molecules and cancer cell environment: challenges for drug development? J. Natl. Cancer Inst. (Bethesda), *91:* 1281–1287, 1999.

59. Killion, J. J., Radinsky, R., and Fidler, I. J. Orthotopic models are necessary to predict therapy of transplantable tumors in mice. Cancer Metastasis Rev., *17:* 279–284, 1999.

60. Kerbel, R. S. What is the optimal rodent model for anti-tumor drug testing? Cancer Metastasis Rev., *17:* 301–304, 1999.

61. Rosenberg, M. P., and Bortner, D. Why transgenic and knockout animal models should be used (for drug efficacy studies in cancer). Cancer Metastasis Rev., *17:* 295–299, 1999.